ENGINEERING ELECTRONICS

BY

DONALD G. FINK

Managing Editor, Electronics

FIRST EDITION

McGRAW-HILL BOOK COMPANY, INC.

NEW YORK AND LONDON

1938

PRINTED IN THE UNITED STATES OF AMERICA

THE MAPLE PRESS COMPANY, YORK, PA.

To

KEITH HENNEY

PREFACE

This book has been written to meet the needs of the practicing engineer who has a good foundation in electricity, but who has no specific training in electronic concepts and methods. The material in the book was originally collected for a lecture course delivered to a group of one hundred such men, members of the engineering department of the Westinghouse Lamp Company. These men, typical of many engineers in the electrical industry, were faced with an increasing number of electronic problems but found that their educational background was sufficiently out-of-date to make the going difficult.

The experience gained in the lecture course has indicated that a practical background in electronic technology should include three basic parts: (1) an understanding of electronic conduction, its capabilities, and limitations, (2) a familiarity with different types of electron tubes and their special fields of use, and (3) knowledge of circuits in which electron tubes are applied. The book has accordingly been divided into three sections: Physical Electronics, Electron Tubes, and Electron-tube Applications.

In writing the book, the author has attempted to steer a course between simple descriptions of equipment on the one hand and elaborate technicalities on the other. The practical side of the subject has been developed in a series of type problems and solutions and by lists of problems at the end of the chapters, numerical answers to which are included in Appendix IV. The mathematical treatment has been limited to high-school algebra and trigonometry, except in two cases where calculus must necessarily be used. These latter demonstrations have been confined to footnotes, where they may be studied or skipped over as the reader desires.

It is hoped that the book will serve not only for the practicing engineer but also for the student who wishes to orient himself in the field before undertaking advanced courses. For such purposes the book is adaptable to an introductory course for

junior or senior college students in general electrical engineering curricula.

The author wishes to acknowledge with gratitude the assistance of Mr. Keith Henney, editor of *Electronics*, who read the manuscript and offered many valuable suggestions, and of Mr. Beverly Dudley, associate editor of the same publication, who spent many hours in reading the proofs. The author has also received many helpful suggestions from engineers of the Westinghouse Lamp Company. Data on the fundamental electron dimensions (e and e/m) were supplied by Professors R. T. Birge of the University of California, J. A. Bearden of Johns Hopkins, and W. V. Houston of the California Institute of Technology.

Illustrations and detailed technical data have been supplied by the following companies: Allis-Chalmers Manufacturing Company, Figs. 1 and 89; General Electric Company, Figs. 77, 78, 81, 84, 87, 88, 91, 92, 97, 104, 105, and 106; RCA Manufacturing Company, Radiotron Division, Figs. 64, 65, 68, 70, 72, 73, 75, 76, 94, 98 and 125; Westinghouse Lamp Division, Westinghouse Electric and Manufacturing Company, Figs. 79, 85, 86, 90, 95, 96, 100, 101, 116, and 117; Western Electric Company, Fig. 93. The cooperation of these organizations is much appreciated.

DONALD G. FINK.

ENGLEWOOD, N. J.,
April, 1938.

CONTENTS

LIST OF TABLES

INTRODUCTION

CHAPTER I

ELECTRONICS IN ENGINEERING—A PRELIMINARY SURVEY

Electronics is that branch of science and technology which relates to the conduction of electricity through gases or in vacuo.

—Proposed Definition, American Institute of Electrical Engineers

Electronic methods in electrical engineering have progressed, in their short history, at a pace which few other branches of applied science can equal. In 1900 no electronic device, save possibly the X-ray tube, was applied to any practical purpose. Today whole industries, including such important activities as radio broadcasting, sound motion pictures, and the world-wide communication systems, are completely dependent upon electron tubes. In research, especially in the physical sciences, electron tubes are now among the most important tools in the exploration of new fields. In industry, electron tubes have solved difficult problems of control and measurement, and, through the use of photosensitive tubes, have made automatic many manufacturing operations which formerly could not be accomplished without the human eye.

Radio receivers, by virtue of their presence in 25 million American homes, have brought into service more electron tubes than there are people in the country. Even more impressive to the imagination is the fact that electronic methods of long-distance communication now make it possible for any one of 34,000,000 telephones in 68 countries to be connected with any other. This record of achievement justifies the claim that electronics has for many years been the most active branch of electrical engineering in producing new and useful applications of electricity.

1. Electronic Functions in Electrical Engineering.—Electrical engineering is concerned with the generation, transmission, and utilization of electric power. In the early history of the art, electric power was generated in batteries, transmitted over short

distances only, and utilized in but the simplest ways. Later it was found that messages could be sent along wires by interrupting the power flow in accordance with a prearranged code. This simple modification of the current flow led to the first large-scale commercial application of electricity, the telegraph. The telephone resulted from the ability of Bell's microphone to modify the current flow in still more complex fashion, producing current variations similar in form to the complicated sound-pressure variations formed by the human voice. In both cases—and in others too numerous to examine—the usefulness of electric energy was expanded by the process of *modifying* the basic current, that is, by producing new forms of current variation capable of performing new functions.

The success of electron tubes is an outstanding example of this principle. Electron tubes have one purpose: to modify the form of the electrical energy which passes through them, according to some prearranged plan. Their utility lies in the fact that they are the most versatile modifiers of electrical energy ever devised.

Rectification.—An important type of power-modification is the conversion of alternating current into direct current, the process called rectification. The alternating current supplied by power companies is unsuited to certain purposes such as charging batteries, electroplating metals, heavy-duty traction service, arc welding, telephone service, and the like, all of which require direct current. The older method of transforming the alternating current into direct current involves two rotating machines, an alternating-current motor and a direct-current generator mechanically coupled together. The electronic method makes use of rectifier tubes which perform the conversion directly. The largest electronic devices in service are the mercury-arc rectifiers (see Fig. 1) which convert alternating current into direct current for subway-train traction motors. They handle a current of five thousand amperes continuously. The inverse process of converting direct current into alternating current is of less practical importance in commercial power practice, but where necessary it may be accomplished in electronic inverter circuits.

Frequency Conversion.—Frequency conversion is a form of power modification which is accomplished almost exclusively by electron tubes. Power, generated commercially at 60 or 25 c.p.s.[1]

[1] For definitions and meanings of abbreviations, see the listing on p. 344.

must be converted for special purposes to alternating current of several thousand cycles per second. Such high-frequency power is used, for example, in the induction furnace, an important tool in alloy research and metallurgical processing. An electron tube, connected in an oscillating circuit, takes direct current or 60-cycle

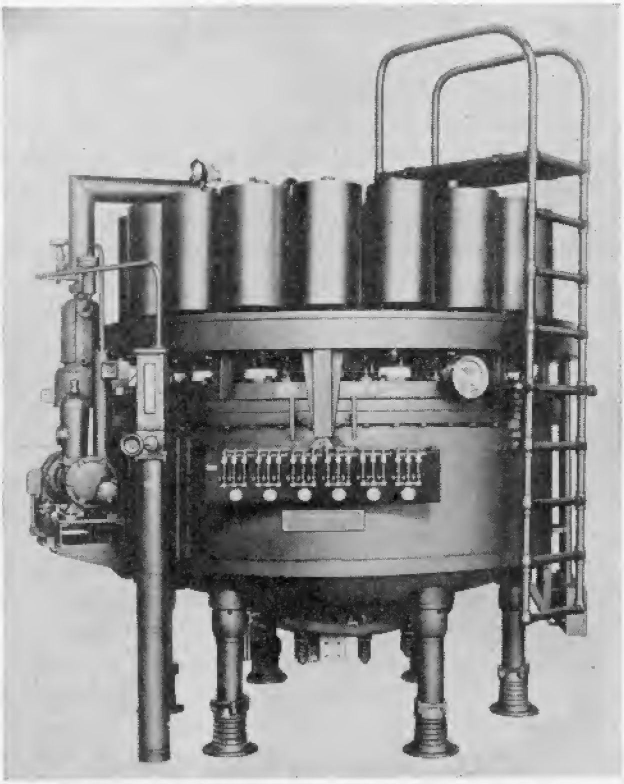


FIG. 1.—Allis-Chalmers tank rectifier, one of the largest electronic devices in service. It is approximately 9 ft. tall and will handle 3000 kw. at 625 volts.

alternating current from the power lines and converts it to alternating current of, say, 5000 c.p.s. While rotating machinery capable of generating currents of this frequency can be built, the electronic oscillator is used in preference because of its greater flexibility and economy.

Alternating current of still higher frequency, from 10,000 to 300,000,000 c.p.s., is essential in radio transmission, since

it is the only type which will radiate from an antenna into space. The generation of these million-per-second currents is a job now assigned exclusively to electron tubes. In a broadcast station,[1] for example, the following electronic functions are utilized: The alternating-current power supplied from the power lines, at 60 cycles, is rectified to direct current by rectifier tubes, converted from direct current into high-frequency alternating current by oscillator tubes, increased in strength by amplifier tubes, further modified in modulator tubes in accordance with the microphone (program) currents, amplified further by very large amplifier tubes, and finally radiated from the antenna.

Amplification.—In addition to alternating- to direct-current conversion and frequency conversion, the electron tube serves to increase the strength of alternating currents without changing their frequencies. This process, called amplification, is the most generally useful of all electronic functions. It consists of taking power from a strong source of supply, say a battery, and modifying its form in a tube operating under the control of a weak source of power. The battery current is thereby transformed into an amplified reproduction of the control current and can be used to actuate devices which the weak control current could not affect.

Repeated Amplification.—In power-engineering practice, the wires over which power is transmitted are made large enough to pass the current without excessive losses of power in heating the wire. In electrical communication, however, it is not economically or technically feasible to make the wires large, because of the great number and length of the circuits. Hence most of the power in the communication current is lost in heating the wire. To prevent the power from being lost entirely, it is amplified at regular intervals along the line. Direct current, supplied locally at each amplification point, is modified by amplifier tubes which operate under the control of the weakened communication current. The modified direct current, thereby given the form of the original communication current, is sent

[1] Tubes in broadcast stations represent a considerable investment. In station WJZ, for example, the tube line-up, exclusive of those in the studio, consists of 2 type 862's ($1650 each) 2 type 892's ($325 each), 3 type 849's ($160 each), 1 type 860, 2 type 865's, 1 type 210, 6 type 857's ($275 each), 10 type 866's, 3 type 217's, 1 type 211, and 10 smaller tubes.

over the next stretch of line. After traveling some distance, it in turn becomes so weak that amplification is again necessary. The amplification is repeated over and over again, new energy being added to the line under the control of the successive "editions" of the communication current. As shown in Fig. 2, in the open-wire transcontinental-telephone circuits, the original voice energy is amplified, in 15 successive stages, a total of 1,000,000,000,000,000 times!

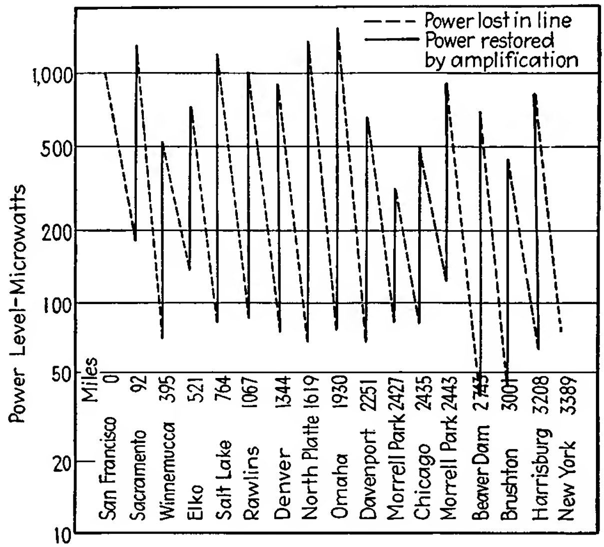


FIG. 2.—Power-level diagram of the transcontinental open-wire telephone circuit between New York and San Francisco, illustrating the use of electronic amplification in communications practice.[1]

Control and Measurement Functions.—In addition to its functions as a rectifier, frequency converter, and amplifier in power and communications practice, the electron tube serves a multitude of uses as a measurement and control device. Its ability to amplify fits it for the detection and measurement of extremely small electric currents and voltages. One form, the electrometer tube, can measure currents of a few hundred electrons per second, the smallest continuous flow of electricity yet detected. Electronic measurement technique is applied generally in all

[1] NANCE and JACOBS, Transmission Features of Transcontinental Telephony, *Trans. A.I.E.E.*, **45**, 1159 (1926).

scientific and engineering pursuits. One of the latest applications is in the field of biology in the measurement of the delicate currents generated by living organisms. Most of our present knowledge of the electrical action of the brain, for example, has been gained through electron-tube technique.

In the field of industrial control, electron tubes find use because of their ability to react to weak stimuli, their rapidity of response, lack of fatigue, and their flexibility of application. An electron tube can respond to a control impulse in less than a millionth of a second, and it can at the same time control many kilowatts of power.

Photosensitive Applications.—In a separate division of the electron-tube family are the phototubes, which are sensitive to light. The eye and the photographic emulsion, the only practical light-sensitive elements available before the advent of the phototube, are limited in many ways. The human eye responds to a limited range of colors, is insensitive, prone to fatigue, and expensive to maintain. Photographic emulsions are severely restricted by the delay and inconvenience of the developing and fixing processes. The phototube, in contrast, is highly sensitive, can be made to have a wide color response, and operates continuously, at any speed, without fatigue. It can detect instantaneously, for example, the light from a star of the fourteenth magnitude which would require several minutes' exposure to record itself on the photographic plate of a large telescope, and which is about $\frac{1}{300}$th as bright as the weakest star visible to the naked eye. In its most advanced form (the iconoscope, described on pages 206–210), the phototube performs all the essential functions of the human eye, perceiving and transmitting complete optical images.

These photosensitive electron tubes have given rise to a wholly new branch of electrical-engineering practice. The applications cover so wide a variety, from sound motion pictures to astronomical research, that it is doubtful if all the phototube uses have ever been listed. A partial list compiled[1] in 1935 contains 253 separate applications of phototubes.

Electronic Sources of Light.—The electronic production of light, while the oldest electronic phenomenon observed (a lightning flash properly falls within the definition of an electronic

[1] *Electronics*, January, 1935, p. 2.

phenomenon, since it is the conduction of electricity through a gas), is one of the latest branches of the field to receive concentrated attention. The mercury arc and the tubular neon lights are familiar examples, as is the newer sodium lamp. The recently developed fluorescent gas-discharge lamps may become serious challengers of the incandescent lamp as a source of general illumination. The high efficiency of the sodium lamp and its effectiveness as a source of street lighting have already given it an important position in the newer installations in that field.

2. The Principles of Electronic Action. The Free Electron and Its Control.—The distinguishing feature of all electron tubes is the fact that they conduct electricity not through wires, as do all other electrical devices, but through a gas or through a vacuum. The utility of this mode of current conduction is the explanation of the utility of the electron tube itself. The following chapters are devoted to the expansion of this basic concept.

Electricity Is a Flow of Electrons.—Wide experimental and theoretical evidence points to the fact that the ordinary electric current is in reality a flow of elementary electrical units, called electrons. These electrons are of unbelievably small size, but they possess a definite mass and other familiar characteristics.

Electrons have the property of repelling one another, the force of repulsion increasing greatly as the electrons near one another. To explain this remarkable behavior, each electron is considered to possess an *electric charge*, the charge being a numerical measure of the force of repulsion experienced between two electrons 1 cm. apart.

The ability of electrons to repel other electrons, *i.e.*, their possession of electric charge, makes it possible to transfer energy through the motion of electrons. When an electric current is generated in a generator, for example, the mechanical energy imparted to the drive shaft is converted into energy of electron motion in the armature winding. Electricity is thus basically the science of electron motion. The conduction of electricity through a wire is simply the motion of electrons along the wire.

With these basic ideas in mind we recall that the electron tube conducts electricity (*i.e.*, conducts a flow of electrons) through a gas or through a vacuum, and that the function of the tube is to

modify the electron flow in some predetermined manner. The action of the electron tube in carrying out this modification process may be shown as follows: Consider a wire carrying a current whose form we wish to modify. To do so we must modify the motion of the electrons flowing along the wire. This can be accomplished only in indirect fashion, so long as the electrons are in the wire. However, if we can remove the electrons from the wire, temporarily, and then return them again, we can modify their motion directly while they are freely moving in space. The temporary removal of the electrons is accomplished by breaking the wire and connecting its two ends to electrodes inside a glass or metal container (or "tube") containing

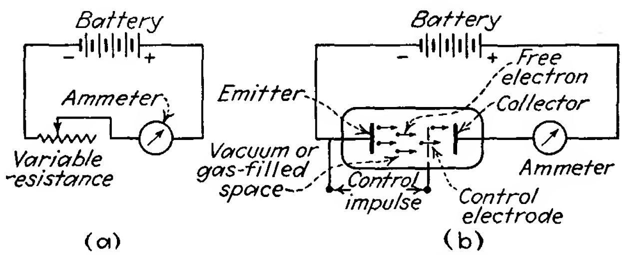


FIG. 3.—In (*a*) the current flowing from the battery may be controlled by varying the amount of resistance in the circuit. The electronic method of control (*b*), by influencing the motion of free electrons, permits more rapid changes of the current, and allows the use of extremely feeble control impulses.

a gaseous or vacuous space. If we cause the current of moving electrons to pass between the electrodes, it is possible to control their motion while they are passing through the tube. This control of the electron motion modifies the electric current flowing in the circuit, and the electronic function is thereby exercised.

It is necessary to satisfy certain obvious requirements in this process. In the first place, the electrons will not leave the first electrode and enter the space beyond without some urging. This urging process, by which electrons are freed from the electrode, is the first essential in any electron tube. In the second place, ways and means must be available for influencing the motion of the electrons while they are in their temporarily free condition before they enter the other electrode. This requirement involves a study of the electric and magnetic forces by which electron motion may be influenced. It requires

study also of the speed with which the electron completes its journey ("transit time") and of the mutual reactions of the electrons on one another during the flight in space ("space charge"). It is also necessary that the electrons be collected by the second electrode and that their subsequent actions be under control; we must inquire what happens to the electrons if they are collected by the second electrode only in part. Finally, the study is incomplete until we examine the influence of the group of gas particles (if they are present in appreciable numbers) through which the electron flow takes place. We find, among other things, that in certain circumstances light is produced by the energy interchanges between the electrons and the gas particles, and that this light can be put to considerable practical use.

3. The Terminology of Electronics. The Organization Chart. The foregoing description of the basic electronic action has been given, for the sake of clarity, with a minimum of technical terms. To introduce the terminology of the subject we refer to the chart shown in Fig. 4, which is intended to serve also as a guide to the organization of this book.

Electronic theory springs from the main body of physics and is particularly dependent on three branches of physical knowledge: electrodynamics, atomic structure, and statistical physics. These give rise to the three basic studies of electronic theory itself: space charge, electron emission, and the gas discharge. Space charge is the study of the motions of groups of electrons in a vacuum of gas-filled space; it reveals the laws which govern the flow of free electrons and the means by which electron motion may be controlled. Electron emission is the study of the "urging process" previously referred to, *i.e.*, the methods by which electrons are made to leave metals and to enter the adjacent space. Three important methods of emission are listed: thermionic emission, caused by heating a metal surface; photoelectric emission, caused by the action of light on a surface; and secondary emission, caused by the bombardment of a surface by electrical charges. The study of the gas discharge reveals the laws governing electron motions in a gas-filled space; the resulting conduction of electricity gives rise to a side branch, the electronic production of light.

The body of electronic theory is then applied to the design and manufacture of electron tubes, of which there are four important

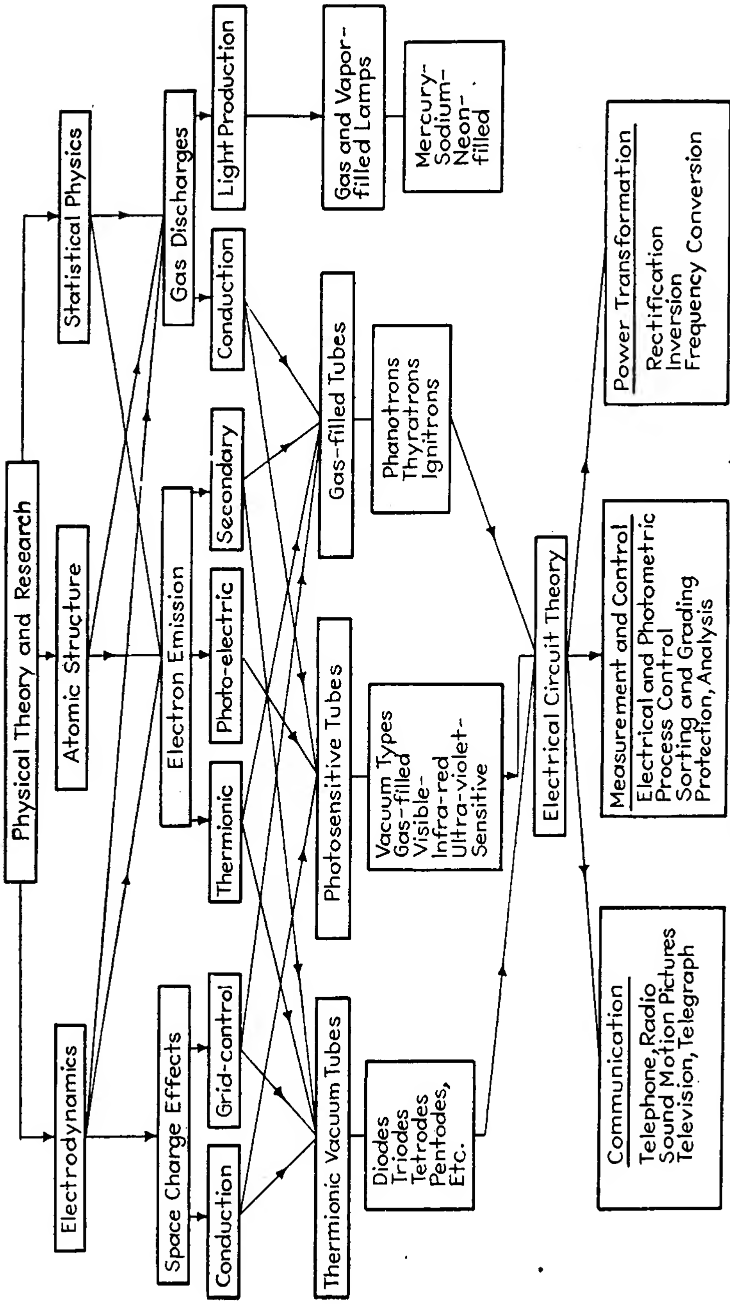


FIG. 4.—Organization of the field of electronics. The arrows indicate the interdependence of the various branches.

types: the thermionic vacuum tubes, photosensitive tubes, gas-filled conduction tubes, and gas-filled lamps. The first two types derive their names from the particular type of electron emission which is employed, but, as is revealed by the maze of crisscrossing arrows, all branches of electronic theory are applied in designing each type. In each type there are a number of special forms, each having different characteristics and applications, some of which are listed.

Electron tubes are of no value, of course, until they are connected with other electrical apparatus. Each different electron-tube application calls for a circuit connection with different types of external apparatus. The study of electron-tube applications begins, therefore, with the study of electrical-circuit theory. Electron-tube circuits are applied, as we have seen, in three major divisions: electrical communication; the control and measurement of scientific and industrial processes; and the transformation of basic power sources.

While the chart is in no sense a complete résumé of electronics, it contains the essential elements, and should serve to guide the reader in his study of the chapters which follow.

PART I

PHYSICAL ELECTRONICS

THE ELECTRON: ITS PROPERTIES, METHODS OF PRODUCTION IN THE FREE STATE, CONTROL BY ELECTRIC AND MAGNETIC FORCES

CHAPTER II

THE FUNDAMENTAL PROPERTIES OF THE ELECTRON

In electronic engineering the properties of the electron are completely basic. While it is possible to design and use electron tubes and electronic equipment without a detailed investigation of the electron itself, an appreciation of its fundamental properties is of the greatest value in understanding the operation of all electronic devices.

4. The Atomic Nature of Electricity.—Since the time of Dalton in the early 1800's, it has been universally recognized that matter consists of separate particles, called atoms, which are the smallest units to which matter, as such, can be reduced. When electricity was first studied, it was believed to exist in the form of a nonmaterial fluid which could be divided without limit into smaller and smaller units. But in Faraday's time it was discovered that electricity is atomic in nature, that is, it consists of definite small units which cannot be further subdivided.

The discovery of the atomic nature of electricity was made through the study of the electrolytic conduction of electricity through liquids. It was found, by passing a direct current through a solution of, say, a silver salt, that metallic silver could be deposited on the negative electrode immersed in the electrolyte. In 1833 Faraday[1] showed that the amount (weight) of silver deposited in a given time was accurately proportional to the strength of the current flow through the electrolyte. In particular it was found that a current equivalent to 1 amp. would deposit 0.001 g. of silver in 1 sec.[2] Since 1 coulomb of electric charge is carried by a current of 1 amp. flowing for 1 sec., there is then 1 coulomb of charge associated with 0.001 g. of silver, or 1000 coulombs with 1 g. of silver. Later it was found that many chemical elements other than

[1] *Phil. Trans.*, 1834, p. 77.

[2] The actual value is 0.001118 g., which is the unit by which the international standard ampere is defined.

silver, both metallic and nonmetallic, exhibited the same behavior and that the amount (weight) of each element deposited in a given time by a given current was always proportional to the atomic weight of the element in question. It was proposed, therefore, that each atom in the solution had a definite electric charge associated with it, and that the charge on each atom was the same.

The number of coulombs of charge associated with each atom can be computed from Avogadro's number, *i.e.*, the number of atoms in a number of grams equal to the molecular weight of the element in question. Using the inaccurate determination of Avogadro's number of his day, G. Johnstone Stoney, in 1874, computed the charge associated with each atom and called this charge the "electron."[1] His value was 10^{-20} coulomb per atom.[2] The modern determination of the value is 1.60×10^{-19} coulomb. Judged by ordinary standards, therefore, the electric charge possessed by a single atom in an electrolyte is extremely small, roughly two-tenths of a billionth of a billionth of a coulomb.

The existence of electrically charged atoms in solutions was well established by 1850, but the presence of free electricity in a gas or in a partial vacuum—on which the science of electronics is based—was not suspected until 1870, when Sir William Crookes[3] performed his famous experiments on the discharge of electricity in gases. Sir William found that if two wires were sealed in a container from which most of the air had been exhausted, a current could be made to pass through the tube if sufficiently high voltage was applied between the wires. When the current passed, a curious fluorescence occurred in the glass container. This fluorescence was shown to be due to the action of rays emanating from the negative electrode (cathode) in the tube. These rays later were proved to be beams of electrons moving at high speed. It was shown that the "cathode rays" could cast shadows of material objects held in their path (see Fig. 5)

[1] In a paper read before the British Association for the Advancement of Science.

[2] Throughout this book it will be necessary to use the exponential designation for very large and very small numbers. The number 10^{20} represents the number 10 multiplied by itself 20 times. The number 10^{-20} is the reciprocal of 10^{20}, *i.e.*, it is the number 1 divided by 10^{20}. If written in full, the number 10^{-20} is 0.000,000,000,000,000,000,01.

[3] *Phil. Trans.*, pt I, 135 (1879).

and that they behaved similarly to light rays in other ways. But their exact nature was not understood for twenty-five years after their discovery.

In 1897 Sir J. J. Thomson determined the speeds at which cathode rays travel, by a very ingenious experiment described fully in Chap. IV (page 68). Similar experiments[1] gave a

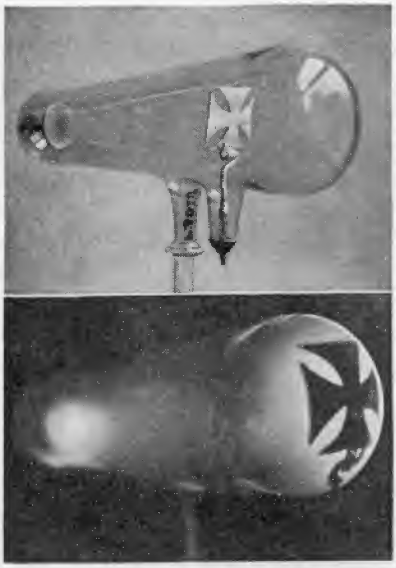


FIG. 5.—Cathode rays. Electrons liberated from the circular electrode are attracted to the cross. Those electrons not intercepted by the cross pass by it and excite fluorescence on the glass container.

value of the mass of each cathode-ray particle (electron), but only in terms of the electric charge associated with it. Assuming that the charge on each particle was the atom of electricity computed by Stoney, *i.e.*, 10^{-20} coulomb, a value for the mass of the cathode-ray particle was found: 0.6×10^{-28} g. Recent determinations give the mass of the electron as 0.91×10^{-27} g.

In a series of brilliant investigations Thomson showed that the cathode-ray particles produced from many different cathode materials possessed the same charge-to-mass ratio. Thomson's work thus went far in establishing the universal presence of the electron in material substances, and showed that the cathode ray

[1] *Phil. Mag.*, **44**, ser. 5, 293 (1897).

was in reality a beam of electrons traveling at high speed. It was established also that the charged particles liberated from hot metals and from metal surfaces illuminated with ultraviolet light had the same value of charge-to-mass ratio as the cathode-ray particle.

In the early 1900's Lord Rutherford established the fact that the negative electricity (the electrons) associated with an atom reside in the outer regions of the atom, and that the positive electricity (protons) reside in an extremely compact and dense core or "nucleus" within the atom. These facts were proved[1] by shooting a beam of positively charged helium atoms (alpha particles from radium) at a thin foil of gold, and noting with a fluoroscope the changes in direction, or scattering, experienced by the atoms in passing through the foil.

In 1913 Bohr suggested that the negative electrons in the atom might be revolving in definite orbits, like planets, about the central positive nucleus. Subsequently he explained the lines in the arc spectrum of hydrogen in terms of this theory with such accuracy that his planetary-electron concept was immediately accepted.

Beginning in 1909, Robert Millikan performed his famous oil-drop experiments, by which he determined the charge on the electron without reference to Avogadro's number. This series of experiments gave the value of the electron charge as 1.59×10^{-19} coulomb, the value which has been used until 1936, when a slight correction was discovered. The value now accepted is 1.60×10^{-19} coulomb. Since the oil-drop experiment is very important in the history of electronics, it is described in detail in Chap. IV (page 55).

5. The Known Properties of the Electron and Their Implications.—According to recent findings of the research physicists, the electron is a negative unit of electricity which may be viewed as an indivisible material particle of the dimensions shown in Table I.

The exact shape of the electron, while commonly assumed to be spherical, is not known. The theory of De Broglie,[2] and the

[1] *Phil. Mag.*, **21**, 669 (1911). See also Chap. VIII in "Radiations from Radioactive Substances," Rutherford, Chadwick and Ellis, The Macmillan Company, New York (1930).

[2] DEBROGLIE, *Ann. d. Physik*, **10**, 322, (1925).

TABLE I.—ELECTRON DIMENSIONS

Mass	0.911×10^{-27} g.
Charge	4.803×10^{-10} electrostatic unit (e.s.u.)
	$= 1.602 \times 10^{-19}$ coulomb.
Diameter	2×10^{-13} cm. approx.

experiments of Davisson and Germer[1] indicate that the electron has a definite wavelike character, but this concept is of small importance in the study of practical electron tubes.

This is a surprisingly small amount of information on so important a unit, but it serves to reveal a great deal of subsidiary information on which the whole electronic art is based.

The Implications of Electronic Mass.—The fact that the mass of the electron is definitely known permits us to make two important types of calculation: (1) It permits the determination of the speed attained by an electron under the influence of a force urging it to move and hence of the time required for the electron to travel from one point to another under the influence of the force. (2) It permits us to determine the kinetic energy (the energy due to motion) possessed by an electron moving at a given speed. Both of these calculations are of interest in the design of practical electron tubes.

When an electron of mass m is acted upon by a force f, it begins to move and gathers speed (accelerates) at a rate which is equal to the force divided by the mass. This is Newton's first law of motion and is expressed as

$$a = \frac{f}{m}, \tag{1}$$

where a is the acceleration (centimeters per second gained each second) attained by the body of mass m, in grams, acted on by the force f, in dynes. If the force is constant and does not change as the position of the electron changes (this condition is assumed in many practical calculations), then the velocity which the body attains at the end of a definite period of time can be calculated, by the expression

$$v = at, \tag{2}$$

where v is the velocity in centimeters per second acquired by the

[1] DAVISSON and GERMER, *Proc. Nat. Acad. Sci.*, **14**, 317, 619 (1928); also THOMSON, *Nature*, **122**, 279 (1928).

electron at the end of t sec. when it is being accelerated from rest at a rate of a centimeters per second per second (cm./sec./sec.). Likewise, if the distance traveled is known, then the time taken to travel the distance can be computed, if the acceleration is known, by the expression

$$t = \sqrt{\frac{2s}{a}} \tag{3}$$

where t is the time in seconds required for an electron, starting from rest, to travel a distance of s cm. when accelerated at a constant rate a cm./sec./sec.

These elementary laws of physics may be applied in a practical example, as follows:

Problem 1. An electron (starting from rest) is acted upon by a force of 9.0×10^{-16} dyne. Find the acceleration of the electron, the time it takes to move 0.5 cm., and the velocity it has acquired when it has moved that distance.

Given:

Mass of electron $m = 0.9 \times 10^{-27}$ g. (from page 21).

Distance of travel $s = 0.5$ cm.

Force acting on electron $f = 9.0 \times 10^{-16}$ dyne.

To find:

Acceleration $= a = \dfrac{f}{m} = \dfrac{9.0 \times 10^{-16}}{0.9 \times 10^{-27}} = 10^{12}$ cm./sec./sec.

Time of flight $= t = \sqrt{\dfrac{2s}{a}} = \sqrt{\dfrac{2(0.5)}{10^{12}}} = 10^{-6}$ sec.

Velocity attained $= v = at = 10^{12} \cdot 10^{-6} = 10^{6}$ cm./sec.

(NOTE.—Those unfamiliar with the method of carrying out multiplication by adding exponents and division by subtracting exponents should review this method in any standard text, since it is a necessary convenience in nearly all electronic calculations.)

The above calculation is based on a practical case. The calculated electron speed (a million centimeters per second) is attained in a distance of 0.5 cm. The implication is, therefore, that an electron can attain a very great speed in a short space.

Kinetic Energy of Electrons in Motion.—When an electron of mass m g. is moving at a velocity of v cm./sec., it then possesses a kinetic energy (K.E.) ergs, given by the expression

$$\text{K.E.} = \tfrac{1}{2}mv^2 \text{ ergs.} \tag{4}$$

This energy is a measure of the ability of the electron to do work. For example, when the electron in Prob. 1 above has moved 0.5 cm., it possesses a kinetic energy equal to

$$\text{K.E.} = \tfrac{1}{2} \times 0.9 \times 10^{-27} \times (10^6)^2 = 4.5 \times 10^{-16} \text{ erg.}$$

This is a very small fraction of an erg, but it is, nevertheless, a great deal of energy for a particle as small as an electron to

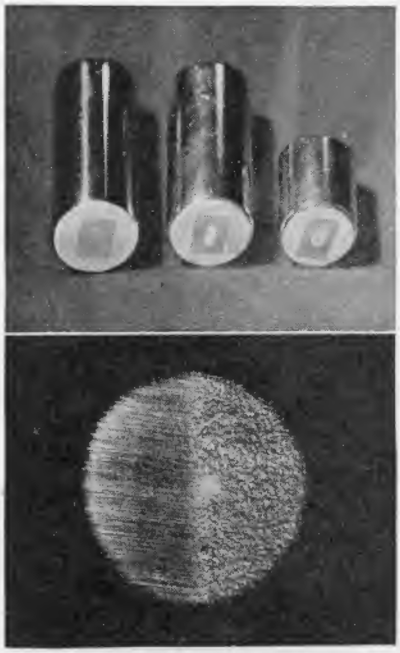


Fig. 6.—Effects of electron bombardment. Above are anodes (tungsten targets mounted in copper) taken from high-voltage X-ray tubes. At the left is a new anode. The etched areas on the others have been caused by the mechanical impact of electrons. Below is a magnified image (50 diameters) of the edge of one of the etched areas. At the left is polished tungsten; at the right the "eroded" surface.

possess. If the electron strikes some object, assuming the electron does not rebound, it delivers its kinetic energy to the obstruction in the form of heat energy. This is the reason why the plates of high-powered vacuum tubes (against which electrons impinge) are hot when operating.

The Implications of Electronic Charge.—It has already been pointed out that the concept of electric charge is employed to

deal with the fact that electrically charged bodies repel or attract one another. If the two charges are alike in sign (both negative or both positive), the force is repelling; otherwise (if the charges are unlike) it is attracting; in either case the strength of the force depends on the distance separating the two charges. If the charge on one body is e_1 e.s.u. and the charge on the other is e_2 e.s.u., then the force of attraction or repulsion is f dynes, when the separation between the charges is s cm., according to

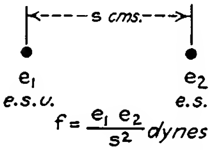


FIG. 7.—The electric force between two "point" charges depends on the product of the charges and the square of the distance between them.

$$f = \frac{e_1 \times e_2}{s^2}. \tag{5}$$

This expression applies, strictly speaking, only to charges small in diameter as compared with the separation between them,

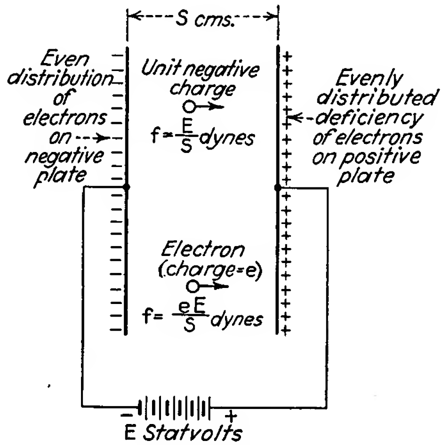


FIG. 8.—The force exerted on a charge located between two large parallel charged plates depends on the voltage applied to the plates and on the distance separating them.

i.e., to "point" charges, as shown in Fig. 7. If the charge is distributed over a surface, as over two conducting plates, then the force between the plates can be calculated only by summing up the individual forces due to all the point charges on the surface. In particular we are interested in two large parallel plates (Fig. 8), since this case is particularly easy to handle and is used as the basis for many practical design problems.

Let us consider two large parallel plates separated by a distance s cm., and with a battery of E statvolts (one statvolt is equal to 300 ordinary volts) connected between them. The action of the battery is to produce an even distribution of electrons on the negative plate and an even deficiency of electrons (even positive-charge distribution) on the other. Now, if we place a unit charge (1 e.s.u.) in the space between these two plates and sum up all the attracting and repelling forces exerted on this unit charge by the two distributions of charge on the plates, it is found that the force f, in dynes, acting on the unit charge is equal to

$$f = \frac{E}{s} \tag{6}$$

Likewise if we place an electron, whose charge is e e.s.u., between the plates, the force f_e acting on the electron, in dynes, is

$$f_e = \frac{eE}{s} \tag{6a}$$

In other words, a battery connected between two parallel plates will set up a force on any electron situated between the plates, urging it to move away from the negative plate and toward the positive. The amount of the force depends on (1) the electronic charge, (2) the voltage applied, and (3) the separation between the plates. This is the principal method employed in electron tubes in causing electrons to move between two electrodes.

It can be shown by calculus that the force acting on a charge between large parallel charged plates is the same whether the charge is near the negative plate, near the positive plate, or anywhere in between. The force does not depend, therefore, on the position of the charge and is consequently a very convenient type of force with which to work, since it permits the use of Equations (2) and (3) without modification.

We may apply Equation (6a) to a typical example, as follows:

Problem 2. Suppose we have two large parallel plates separated by 0.5 cm., and assume that a battery of 1 statvolt (300 volts) is connected to them. Find the force on an electron situated between the plates, the acceleration which results from the force, the time it takes the electron to go from the first plate to the second, starting from rest, and the velocity it attains on reaching the second plate:

Given:

$E = 1$ statvolt.

$s = 0.5$ cm.

$$\left.\begin{aligned} e &= 4.8 \times 10^{-10} \text{ e.s.u.} \\ m &= 0.9 \times 10^{-27} \text{ g.} \end{aligned}\right\} \text{(from page 21)}$$

To find:

$$f = \frac{eE}{s} = \frac{(4.8 \times 10^{-10})(1)}{0.5} = 9.6 \times 10^{-10} \text{ dyne}$$

$$a = \frac{f}{m} = \frac{9.6 \times 10^{-10}}{0.9 \times 10^{-27}} = 1.07 \times 10^{18} \text{ cm./sec./sec.}$$

$$t = \sqrt{\frac{2s}{a}} = \sqrt{\frac{2(0.5)}{1.07 \times 10^{18}}} = 0.97 \times 10^{-8} \text{ sec.}$$

$$v = at = (1.07 \times 10^{18})(0.97 \times 10^{-9}) = 1.04 \times 10^{9} \text{ cm./sec.}$$

This velocity v is enormous,[1] about 1/30th the speed of light. The example shows that a 300-volt battery connected between plates separated about 1/5 in. (about 0.5 cm.) can cause an electron to move with such great speed that it will cover the distance between the plates in about a billionth of a second. Typical electron speeds for different accelerating fields and distances are given in Fig. 10.

The Implications of the Electronic Charge-to-mass Ratio.—If we calculate the value of the ratio of the electron's charge to its mass, we obtain the following value:

$$\frac{e}{m} = \frac{4.8 \times 10^{-10}}{0.9 \times 10^{-27}} = 5.3 \times 10^{17} \text{ e.s.u./g.}$$

The charge on the electron is thus enormous compared with its mass; a gram of electrons has associated with it nearly a billion, billion e.s.u. of charge. This fact of nature gives rise to the fundamental ability of the electron to move rapidly under the repelling influence of near-by electrons. To show the enormous force associated with a moderate electronic mass,

[1] When speeds as great as this are encountered, the mass of the moving electron begins to increase appreciably. The new mass m_v at velocity v cm./sec. is given in terms of the original mass (at rest) m by the Lorentz-Einstein equation:

$$m_v = \frac{m}{\sqrt{1 - (v^2/c^2)}},$$

where c is the speed of light (3.0×10^{10} cm./sec.). In the above example the increase in mass is less than 1 per cent. This "relativity" increase in mass is not important in any but the more advanced electronic computations.

imagine two spheres made up entirely of electrons, weighing 1 g. each (about 1/28 oz.) and separated from one another by 100 km. (62 miles, 10^7 cm.). According to Equation (5), these

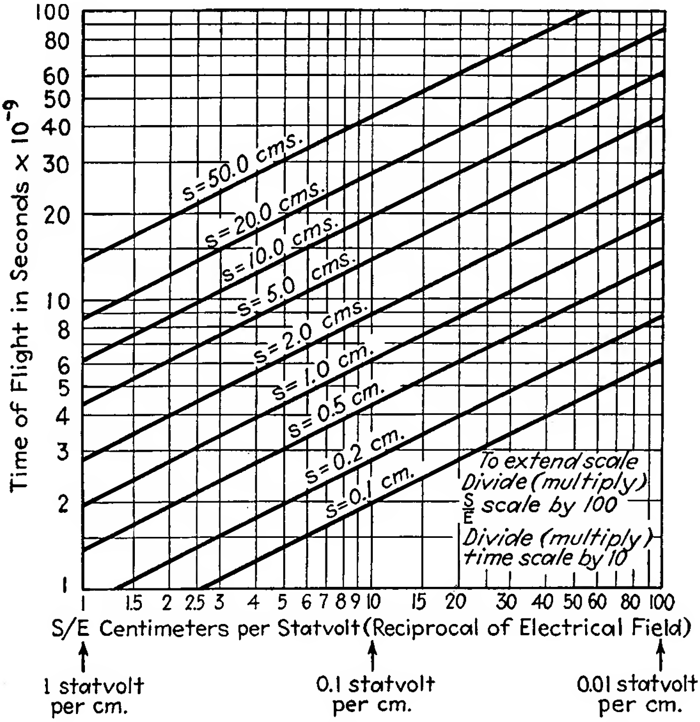


FIG. 9.—Duration of electron flight between large parallel plates in terms of applied voltage E and distance s separating the plates. Calculated by the method of Prob. 2.

two spheres of charge will repel each other with a force f:

$$f = \frac{e_1 \times e_2}{s^2} = \frac{(5.3 \times 10^{17})^2}{(10^7)^2} = 2.8 \times 10^{21} \text{ dynes.}$$

Since there are 444,823 dynes to the pound and 2000 lb. to the ton, this repelling force is equal to three million, million tons! This despite the fact that the spheres weigh only 1 g. and are separated some 60 miles.

The high value of the charge-to-mass ratio and the consequent ability of the electron to exert a force very great compared

with its mass are the fundamental properties on which the utility of electronic conduction of electricity depends. This fact explains the enormous speeds which electrons can attain in a short space and under the influence of moderate voltages. It implies also that the electron is extremely agile. Not only can it attain high speeds quickly, but it can come to a stop with equal

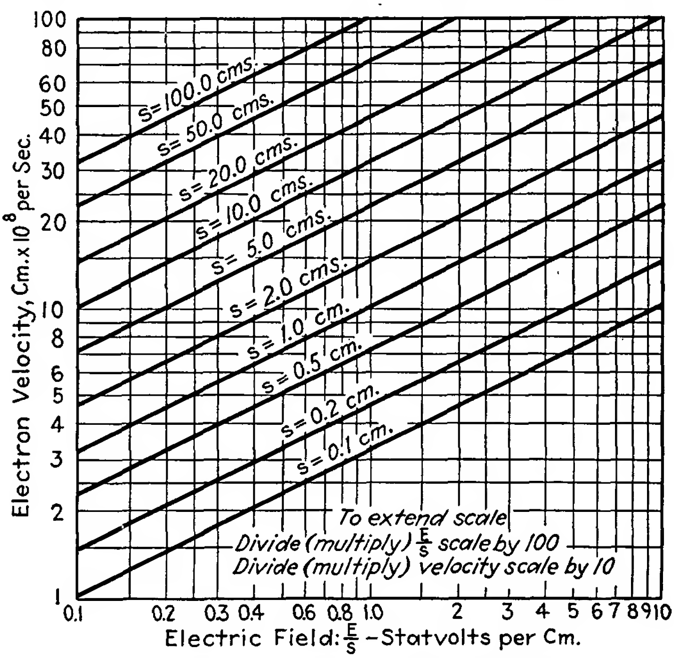


FIG. 10.—Electron velocity, in terms of applied electric field and distance of travel. The lowest value of speed shown is 100 million centimeters per second!

suddenness, reverse itself, and speed up as quickly in the opposite direction, all within the duration of a fraction of a millionth of a second. This flexibility of electron motion is utilized in producing the complex and rapidly varying currents used in electrical communication and in the rapid control of industrial equipment and processes.

6. Electrons in Groups. Energy Transfer. The Electron in Current Flow.—In the preceding sections the properties of the individual electron have been examined. In practical applications, except in the most advanced phases of electronic research, we never deal with single electrons, not only because they are

difficult to isolate but also because the effect of a single electron is so very small. The practical unit of electric charge, the coulomb, is composed of an enormous number (0.624×10^{19}) of electrons. With this fact in mind, it may be inquired why the motion of an individual electron is of any consequence. The answer is obvious: If we can describe the motion of a single electron, we can with suitable modifications describe the motion of any number of them. The calculation of the energy amassed by a single electron in passing between two parallel plates is thus a clue to the energy transfer associated with the motion of large groups of electrons.

The reader is familiar with the fact that when a steady current flow of 1 amp. flows in a circuit and when the voltage in the circuit is 1 volt, then 1 watt of power is being generated or dissipated in that circuit. The watt is the unit of the rate of energy transfer; it is equivalent to the transfer of 10^7 ergs in 1 sec. The watt-second, the practical unit of electrical energy is, therefore, 10^7 ergs. We may now inquire the meaning of these units of power and energy transfer in terms of the electron motion involved in electric-current flow.

Consider again the parallel plates separated by s cm. between which is connected a battery of E statvolts. A single electron (charge e) between the plates experiences a force urging it toward the positive plate, the force being eE/s dynes. Now, if the electron starts from the negative plate and moves toward the positive plate, the energy W it acquires, in ergs, is the product of the force times the distance moved. The equation is

$$W = f \times s = \frac{eE}{s} \times s = eE. \tag{7}$$

The energy associated with electron motion is thus the product of the electron charge e and the voltage difference E through which it moves. If we are concerned with a large number n of electrons, then the energy is n times as great.

Suppose, for instance, that the number of electrons involved is $n = 0.624 \times 10^{19}$ (1 coulomb). If this amount of charge moves between two plates, across which there is a voltage difference of 1 volt ($1/300$ statvolt), the total energy transfer W_t is then

$$\begin{aligned} W_t &= neE = (0.624 \times 10^{19})(4.8 \times 10^{-10})1/300 \text{ ergs} \\ &= 10^7 \text{ ergs.} \end{aligned}$$

The unit of energy associated with the motion of 1 coulomb of electrons through a potential difference of 1 volt (1 volt-coulomb or 1 watt-second) is thus 10^7 ergs.

Power.—If the 1 coulomb of charge considered above moves past any given point in the circuit in 1 sec., then the rate of flow of charge is 1 coulomb per second, or a current flow of 1 amp. During each second in the case cited 1 coulomb of charge moves through the voltage difference of 1 volt, and the rate of energy transfer is 10^7 ergs per second, or 1 watt of power.

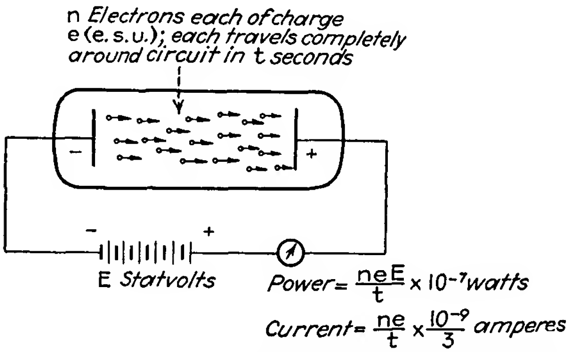


FIG. 11.—Power and current flow in a circuit involving free electrons.

Power, the rate of energy flow, is thus determined by the amount of charge (number of electrons), the velocity at which it moves, and the voltage difference through which it passes. If the voltage difference is 1 volt, then a power flow of 1 watt represents 0.624×10^{19} electrons flowing past each point in the circuit every second. These current and power relationships are shown in Fig. 11.

Not all types of current flow are exclusively composed of electrons. In metallic circuits it is generally agreed that the current consists of free electrons, but in conduction through liquids or gases part of the current is carried by positive charges. When both positive and negative charges are present (Fig. 12), the motion considered is the *relative* motion of the negative charges past the positive ones. So far as the external circuit is concerned, it makes no difference whether the current is carried by negative charges or by positive charges, or by both, so long as all measuring and power-dissipating equipment responds equally well to both types of charge motion, which is generally the case.

7. Magnetic Effects of Electron Motion.—When two electric charges move relative to each other, a peculiar type of mutual reaction occurs which is described in terms of the magnetic field. Consider a long straight wire, along which electrons are flowing. Around this wire exists a magnetic field resulting from the current flow. An electron in motion outside the wire, immersed in this magnetic field, will experience a force, the magnitude of that force depending on the direction of motion of the external electron with respect to that of the electron current in the wire, and on the distance between the wire and the electron. The direction of the force on the electron is always at right angles to its motion, so the effect of the magnetic field is to deflect the electron from its original line of motion.

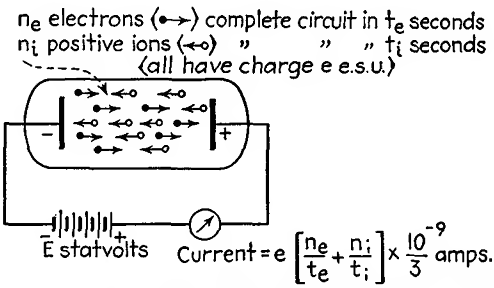


FIG. 12.—Current flow when both positive and negative charges are present.

In practical cases the magnetic field is made nearly uniform (that around a straight wire is not uniform), and in this case the direction of the electron taken under the influence of the magnetic field is not difficult to calculate. The motion is usually circular or in spiral form, since the force is at right angles to the electron motion at any instant. In Chap. IV some details of the magnetic control of electron motion are discussed.

Problems

1. Calculate the acceleration undergone by an electron under the influence of a force equal to 10^{-9} dyne.

2. A 200-volt battery and two large metal plates are available. If the plates are parallel, calculate the distance at which they must be placed apart to exert a force of 10^{-9} dyne on an electron between them.

3. An electron, starting from rest at a negative plate moves to a positive plate parallel to the first and 0.3 cm. distant from it. The battery between the plates has a strength of 200 volts. How fast does the electron move

when it reaches the second plate, and how long is it in flight between the plates?

4. What is the kinetic energy of the electron in Prob. 3, when it reaches the positive plate? What part of the circuit supplies this energy? To what part is it delivered?

5. What is the *average* speed of the electron in Prob. 3? What is the average strength of the current flow, in amperes, of this single electron during its flight? How many electrons, under the same conditions and assuming no mutual reaction between electrons, would make up a current of 1 amp.?

6. How many electrons flow through the filament of a 60-watt incandescent lamp in one year when the applied voltage is 115 volts? What is the combined mass of these electrons? How does this mass compare with the weight of the filament itself (0.015 g.)? How long must the current flow before the combined mass of all the electrons participating in the current flow is equal to the weight of the filament itself?

7. Taking into account the Lorentz-Einstein increase in mass (see footnote, page 26), calculate the approximate voltage necessary to accelerate an electron to a speed equal to 95 per cent the speed of light (3×10^{10} cm./sec.). The kinetic energy of an election moving at speed v is

$$W = mc^2 \left\{ \frac{1}{\sqrt{1 - v^2/c^2}} - 1 \right\} \text{ ergs}$$

where c is the speed of light and m is the mass of the electron at rest.

CHAPTER III

EMISSION. THE PRODUCTION OF FREE ELECTRONS

Introduction.—There are three elements in electronic conduction: the release of electrons from one electrode, the control of their motion while they are free in space, and their subsequent capture by another electrode. The first of these three processes, by which the electrons are released from the first electrode, is called *electron emission.* The principal methods of obtaining free electrons by emission and a brief outline of the theory underlying each method are presented in the following sections.

8. Electrons in Metals.—Electrons exist in nature principally in the outer region of each atom of matter, being held in that vicinity by a balance between the attractive force exerted by the positive nucleus of the atom, on the one hand, and the centrifugal force of rotation, on the other. When atoms form crystalline solids, as in solid metals, the atoms arrange themselves in an orderly latticelike structure whose shape is characteristic of the crystal itself. The forces which maintain the lattice arrangement are not available for maintaining the electrons in their positions. Consequently a small percentage of the electrons associated with each atom find themselves free to move. The degree of motion is restricted in insulators, but in metallic conductors some of the electrons become *unbound* and are free to move from atom to atom. These unbound electrons are the agents by which electric current is carried through conductors, the current being simply an orderly motion of the unbound electrons under the influence of electric or magnetic forces. In the absence of such forces, however, the unbound electrons assume a more or less chaotic motion, moving with a wide range of speeds and in every possible direction. This is the state of affairs from which electron emission is to be obtained.

It is found that, if the electrons move toward the surface of the metal with sufficient velocity, they will break through the surface and become, at least momentarily, free electrons. To secure

this "breaking through," it is necessary to impart energy to the electrons from some external source, since the ordinary energy of the unbound electron is small. This external energy may come from a variety of sources; it may come from heat energy, from the energy stored in electric or magnetic fields, from the energy of light, or from the kinetic energy of electric charges bombarding the metal surface.

The Types of Emission.—In order of practical importance, the four principal methods of obtaining free electrons from solid metals are:

1. Thermionic emission, in which the releasing energy is secured by heating the metal and thus increasing the thermal energy of the unbound electrons.
2. Photoelectric emission, in which the energy of light falling on the metal surface is transferred to the unbound electrons.
3. Field emission ("cold emission"), in which the electron energy is increased by the presence of an electric field at the surface of the metal, *i.e.*, by a concentration of positive charges external to the surface.
4. Secondary emission, in which the necessary energy for release is obtained from the kinetic energy of electric charges bombarding the metal surface.

In addition to these methods of obtaining emission from solids, ionization should be mentioned as the principal method of obtaining free electrons from gases and vapors, which is of great practical importance. Ionization is discussed in Chap. V.

Work Function, the "Energy of Surface Restraint."—The amount of energy required to remove an electron from a metal surface is known as the *work function* of that particular metal. By various means, examined in later sections, the work-function values for various metals have been measured. The energy required varies roughly from 10^{-12} to 10^{-11} erg for the various metals and materials practically used.

For convenience these values of energy are not expressed in ergs but in terms of the voltage difference which gives the electron the same value of energy. Equation (7) on page 29 states that when an electron of charge e e.s.u. moves through a voltage difference of E statvolts, the energy it acquires is

$$W = eE \text{ ergs} \tag{7}$$

or, if we express the voltage difference in volts, the expression is

$$W = \frac{eE'}{300} \text{ ergs,} \tag{7a}$$

where E' is the voltage difference expressed in volts. This expression may be rearranged to give the voltage E' which imparts to the electron a given energy W:

$$E' = \frac{300W}{e} \text{ volts} \tag{8}$$

Substituting the values of energy given above, and the charge (in e.s.u.) on the electron, we find that the work functions of

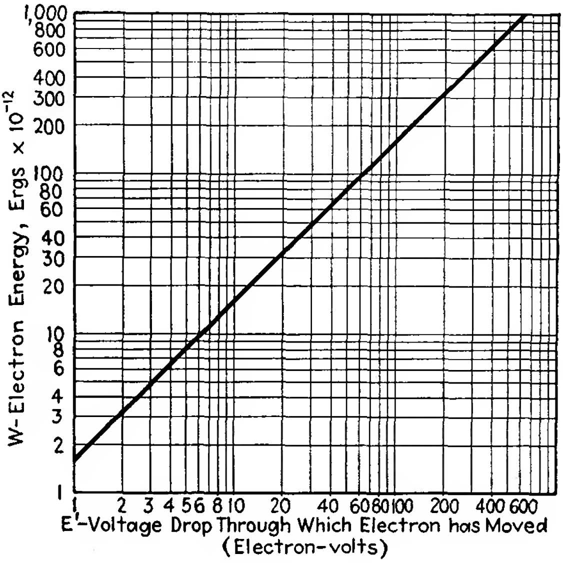


FIG. 13.—Conversion chart, electron-volts to ergs, computed with the aid of Equation (8).

practical electron emitters, expressed in equivalent volts, range from $E' = 300 \times 10^{-12}/4.8 \times 10^{-10} = 0.625$ volt to

$$E' = \frac{300 \times 10^{-11}}{4.8 \times 10^{-10}} = 6.25 \text{ volts.}$$

The virtue of expressing the work function in volts is twofold: It provides numbers easy to remember, and it expresses the fact that voltage differences are used in electron tubes to endow electrons with energy. Table II gives the work-function-voltage (ϕ_0) values for various metals and metallic oxides which are used as practical electron emitters.

The Nature of the Surface Restraint.—It may be wondered what forces are at work at the surface of the metal to restrain the electrons. The force is very complex, but its most important

TABLE II.—CONSTANTS OF THERMIONIC EMITTING SUBSTANCES

Substance	A, amp./cm²/deg.²	ϕ_0, volts	$b_0 = \frac{\phi_0}{k}$, °K.
Tungsten (W)	60	4.52	52,400
Thoriated tungsten (Th-W)	3	2.63	30,500
Barium-strontium oxides (BaO-SrO)	10^{-2}	1.04	12,000
Molybdenum (Mo)	55	4.15	48,100
Platinum (Pt)	17,000	6.26	72,500
Nickel (Ni)	1380	5.03	58,300
Tantalum (Ta)	60	4.07	47,200

constituent is the simple electrical-attraction force which exists between charges of unlike sign. When an electron having a

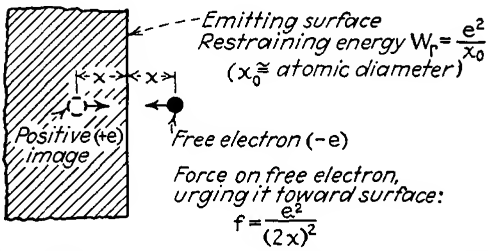


FIG. 14.—Mirror image induced in a metal surface by the removal of an electron, resulting in a restraining force on the electron.

negative charge of 4.8×10^{-10} e.s.u. leaves the surface of a metal, the metal thereby becomes charged with a *positive* charge of the same value. The removal of the electron, in other words, "induces" a countercharge in the surface. The apparent position of this countercharge is behind the surface by the same amount that the electron is in front of the surface, *i.e.*, the positive charge is a "mirror image" of the electron, as shown in Fig. 14. Since this is true, the force acting to pull the electron back to the surface may be found by Equation (5) (Chap. II, page 24), $f = e_1e_2/s^2$, where s, the separation between the charges, is equal to twice the distance between the electron and the surface. To find the energy required to remove the electron, we have only

to multiply this force by the distance moved against it, taking into account the fact that the force decreases as the electron travels away from the surface. This multiplication can be performed by the methods of calculus, with the following result:

$$W_r = \frac{e^2}{4x_1} \text{ ergs,} \tag{9}$$

where W_r is the energy necessary for removal against the "image force," e is the charge on the electron, and x_1 is the separation between the surface and the electron when it begins to move.[1] Choosing the proper values of x_1 is difficult since, if the electron starts to move directly at the surface, x_1 is zero and the required energy W_r is infinite. But if we determine the value of W_r experimentally, it is found that the expression which fits the facts is

$$W_r = \frac{e^2}{x_0} \text{ ergs,} \tag{9a}$$

where x_0 is a distance about equal to the diameter of the metal atoms which compose the surface. This fact indicates that the metals whose atoms have large values of diameter should have small values of work function, and hence should emit electrons most readily. This is found to be true. The alkali metals, rubidium, caesium, barium, and strontium, all of which have low work functions, also have large atomic diameters. The energy of surface restraint, or work function, of a metal is not a simple electrical image force, but the expression in Equation (9a)

[1] The derivation of this equation follows:

$$\begin{aligned} W_r &= \text{force} \times \text{distance} = \int f\,ds = \int \frac{e^2}{s^2}\,ds \\ &= \int_{x_2}^{x_1} \frac{e^2}{(2x)^2}\,dx \\ &= \left(\frac{e^2}{4}\right)\left[-\frac{1}{x}\right]_{x_2}^{x_1} = \left(\frac{e^2}{4}\right)\left(\frac{1}{x_2} - \frac{1}{x_1}\right). \end{aligned}$$

Now if the distance x_2 to which the electron is removed is large (as it is in any case), the term $1/x_2$ may be neglected, and the expression becomes

$$W_r = -\frac{e^2}{4x_1}$$

which is that given above. The negative sign indicates that the work is performed *on*, rather than *by*, the electron.

gives a good approximation to the actual force, including its electrical as well as nonelectrical components.

9. Thermionic Emission.—The fact that heated metals emit electrons much more readily than do cold was known in the late 1800's, but it did not receive a systematic investigation until 1901, when O. W. Richardson began an analysis which remains, with some modifications, as the basic theory underlying thermionic emission.[1] Richardson's work was based on thermodynamics, and the equation he developed contained thermodynamic constants. From 1918 to 1922 Laue[2] and Dushman[3] independently derived the same equation from quantum-theory considerations and determined the constants of the equation from the electronic properties and quantum conditions involved. Later work by Fermi and Dirac established a theory of electrons[4] in metals which differs considerably from that assumed by Richardson, but the Richardson equation still stands as the practical embodiment of thermionic theory.

The basis of the theory is the fact that the unbound electrons in the metal surface have a random motion, having a wide range of speed, moving in any possible direction. When the metal is heated, the thermal energy adds to the vigor of this random motion, *i.e.*, the speed of the unbound electrons increases rapidly as the temperature increases.[5] When this speed increases to the point where the kinetic energy of an electron is in excess of the restraining energy of the work function, then the electron becomes free and is emitted from the surface. Let the restraining energy (work function) be W_r and the speed of the electron, *in the direction of the surface*, be v. Then the kinetic energy of the electron (mass = m) due to this motion is $\frac{1}{2}mv^2$, and this kinetic energy must be equal to or greater than the restraining energy, *i.e.*,

$$\tfrac{1}{2}mv^2 \geqq W_r \tag{4a}$$

[1] For a full account of this work see, O. W. Richardson, "Emission of Electricity from Hot Bodies," Longmans, Green & Company, New York, 1921. See also *Phil. Mag.*, **28**, 633 (1914).

[2] Laue, *Jahrb. d. Rad. u. Elektronik*, **15**, 205, 257, 301 (1918).

[3] Dushman, *Phys. Rev.*, **21**, 623 (1923).

[4] Sommerfeld, *Zeits. Physik.*, **47**, 1 (1928).

[5] According to the Fermi-Dirac theory, only the higher energy electrons experience this increase in energy from the added thermal energy, but the effect is the same.

or, rearranging,

$$v \geqq \sqrt{\frac{2W_r}{m}} \text{ cm./sec.} \tag{10}$$

Richardson's task was to discover how many electrons would have a speed equal to or greater than $\sqrt{2W_r/m}$ when the metal was heated to a given temperature. This task was a highly involved one; it made use of the Maxwell-Boltzmann law giving the speed distribution of electrons, which is rooted in probability theory. The final result is the Richardson equation, which has stood the test of use during the entire history of electronic discovery.

The Richardson Equation.—We consider a pure-metal surface heated to a temperature T °K. (273° plus the temperature of the metal in degrees Centigrade), whose work function is ϕ_0 volts. We wish to know how many electrons will be emitted in 1 sec., from 1 sq. cm. of the metal surface. For practical purposes we state the emission in amperes (an ampere is, we remember, 0.624×10^{19} electrons per second), and give the emission the symbol I_s:

$$I_s = AT^2\epsilon^{-\frac{\phi_0}{kT}} \text{ amp./sq. cm.} \tag{11}$$

A is a constant which has a theoretical value of 60.2 for all pure metals, and k is Boltzmann's constant, expressed as 0.863×10^{-4} volts per degree. This equation is as difficult to evaluate mentally as it is useful in practical electronic problems. It consists of three separate factors, which are best examined separately.

The first factor is the constant A; its theoretical value for pure metals, as revealed by the work of Laue and Dushman, is 60.2, a figure in fairly close agreement with the values found in experiments. For impure metals the value of A is usually much higher; in platinum (with which oxygen is always associated as an impurity) the experimental value of A is 17,000. Values of A for various practical emitters are tabulated in Table II.

The second factor in the equation is T^2, the temperature in degrees Kelvin, squared. The effect of this factor is to increase the number of electrons emitted as the temperature increases. So far as *this factor* is concerned, the emission quadruples when the temperature is doubled, increases ninefold when the tempera-

ture is tripled, and so on. Actually the emission may increase by several million times when the temperature is doubled, because the effect of the third factor is much greater than that of the second.

The third factor $\epsilon^{-\frac{\phi_0}{kT}}$, is the number $\epsilon = 2.718$ raised to the $-\phi_0/kT$ power. To evaluate the effect of this factor, we may choose some particular metal, say, pure tungsten, and calculate the value of the factor at various temperatures. The value of ϕ_0 for pure tungsten is 4.52 volts. At a temperature of $T = 1600°$ K. (slightly above red heat), the value of the exponent is

$$\frac{\phi_0}{kT} = \frac{4.52}{0.863 \times 10^{-4} \times 1600} = 32.8.$$

The value of the factor is then $\epsilon^{-32.8}$. When the temperature is changed, to 3300° K. (somewhat below the melting point of tungsten), the exponent is changed, to a value of 15.8. This decrease in the exponent means an enormous increase in the value of the factor $\epsilon^{-\frac{\phi_0}{kT}}$. The factor increases by $\epsilon^{32.8-15.8}$ times, which is approximately 24,000,000 times. Comparatively small changes in temperature thereby produce relatively enormous changes in the electron emission. In pure tungsten the emission is about a millionth of an ampere per square centimeter at 1600° K., and about a hundred amperes per square centimeter at 3300°K.

For the same reason, relatively small changes in the work function of the material used can produce similarly enormous effects on the emission. Halving the work function has exactly the same effect as doubling the temperature.

Substances Used for Obtaining Thermionic Emission.—The Richardson equation reveals that two conditions must be met in a practical thermionic emitting substance: low work function and the ability to operate at high temperatures. It is necessary to select metals and metallic compounds for emitting purposes with these factors in mind. The materials used are divided in two classes, those with fairly high work functions, which work at very high temperatures, and those with lower work functions, which supply usable amounts of electron emission at lower temperatures. In the first group are the pure metals tungsten, molybdenum, and tantalum, all of which have high melting

points. In the second group are thoriated tungsten (a combination of a thorium layer on a tungsten base, which has a low work function but must be operated at low temperature to preserve the thorium layer), and the metallic oxides of barium and strontium, which have very low work functions and can be operated readily at low temperatures.

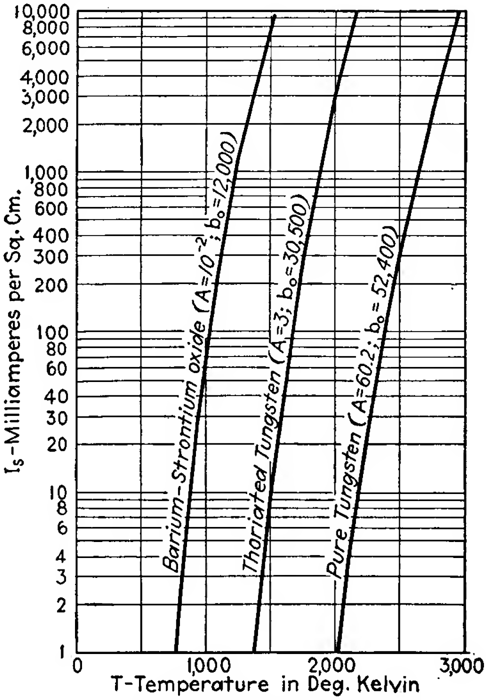


FIG. 15.—Emission currents of three practical emitters, calculated with the aid of Equation (11).

Typical emission values, in amperes per square centimeter, at various temperatures, of several substances are shown in Fig. 15. Table II gives the material constants A, ϕ_0, and the quantity $b_0 = \phi_0/k$ for each of these substances. In practical electron tubes, those substances used to the greatest extent are pure tungsten, and barium-strontium oxide, the former in large tubes where high voltages are present and the latter in smaller tubes, such as those used in radio receivers, where the voltages applied or not higher than several hundred volts. Thoriated tungsten, once widely used as an emitting substance in low-voltage radio

receiving tubes, has now been almost completely replaced by barium-strontium oxide, but it is used in low- and medium-powered transmitting tubes. An interesting property of thoriated tungsten as an electron emitter is its ability to renew itself under proper heat treatment. The thorium layer is replenished by heating the filament to a temperature considerably above operating temperature for a short time.

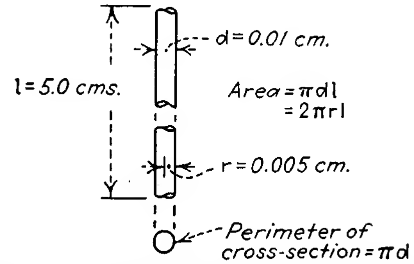


FIG. 16.—The surface area of a cylindrical filament is equal to the perimeter of its cross section times its length (Prob. 1).

To illustrate the use of Richardson's equation, the following problems should be considered:

Problem 1. Calculate the total emission, in amperes, available from a pure tungsten wire, 5 cm. long and 0.01 cm. in diameter, operated at a temperature of 2500°K.

Given:

$A = 60.2.$
$T = 2500°\text{K}.$
$\epsilon = 2.718.$
$b_0 = 52{,}400 \left(= \frac{\phi_0}{k} \right)$
$l = 5 \text{ cm}.$
$d = 0.01 \text{ cm}.$

To find:

$$
\begin{aligned}
I_s &= AT^2\epsilon^{\frac{-b_0}{T}} \text{ amp./sq. cm.} \\
&= 60.2 \times (2500)^2 \times (2.718)^{\frac{-52{,}400}{2500}} \\
&= 3.76 \times 10^8 \times (2.718)^{-20.9} \\
&= 3.76 \times 10^8 \times 0.795 \times 10^{-9} \\
&= 0.299 \text{ amp./sq. cm.}
\end{aligned}
$$

The area of the surface of the wire (see Fig. 16) is

$$\pi\, dl = 3.1416 \times 0.01 \times 5 = 0.157 \text{ sq. cm.}$$

The emission from the surface is therefore

$$0.299 \times 0.157 = 0.047 \text{ amp., total.}$$

Problem 2. Calculate the total emission available from a barium-strontium surface whose area is 0.157 sq. cm. at a temperature of 1100°K.

Given:

$A = 10^{-2}$

$T = 1100°K.$

$\epsilon = 2.718.$

$b_0 = 12{,}000.$

$\text{Area} = 0.157.$

To find:

$$I_s = AT^2\epsilon^{\frac{-b_0}{T}}$$

$$= 10^{-2}(1100)^2 2.718^{\frac{-12{,}000}{1100}}$$

$$= 1.21 \times 10^4 \times 2.718^{-10.9}$$

$$= 1.21 \times 1.85 \times 10^{-1} = 0.224 \text{ amp./sq. cm.}$$

$$I = I_s(\text{Area}) = 0.224 \times .157 = 0.0352 \text{ amp.}$$

It will be noted that the emission for the barium-strontium surface, per square centimeter, is roughly equal to that of the pure tungsten, despite the fact that its operating temperature is 1400°K. lower than that of the tungsten.

Practical constructions used in thermionic emitters, and the methods of preparing the emitting substances, are discussed under "Thermionic Cathode Structures," Chap. VI, page 96.

10. Photoelectric Emission.—The second important method of obtaining electron emission from metals is photoelectric emission, in which the release energy is transferred to the electrons within a metal directly from light falling on the metal surface. This effect, on which so many important electronic applications depend, was first discovered in 1887 by Hertz,[1] who noticed that a spark would jump between two electrodes much more readily if the negative electrode was illuminated with strong ultraviolet light. The photoelectric effect was studied systematically by Hallwachs, but it was not until 1905 that Einstein gave a satisfactory explanation of its behavior.

This explanation depends on the quantum theory of light, in which is assumed that a beam of light consists of discrete "bundles" of energy, or *quanta*, the amount of energy in each quantum being greater the higher the frequency of the light. The energy of a light quantum falling on a metal surface may be

[1] This work was in connection with Hertz's discovery of radio waves. See *Ann. d. Physik*, **36**, 769 (1889).

transferred directly to an unbound electron within the metal, thereby giving it sufficient energy to overcome the surface forces. The more powerful the quantum (the more energy it contains), the more readily will the electron be released. Since the quantum energy increases as the color progresses from the low-frequency infrared through the visible colors to the high-frequency ultraviolet, it is found that ultraviolet light is highly effective in producing photoelectric emission. This is true especially when the light strikes metal surfaces whose work function is high, and which require more energy per electron to secure emission.

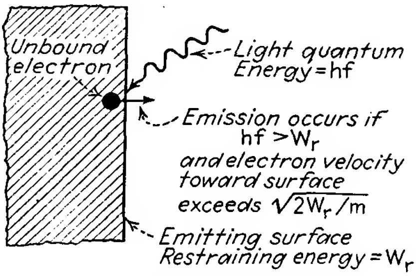


FIG. 17.—Photoelectric emission.

To produce large amounts of photoelectric emission with light of lower frequencies, *i.e.*, the red end of the spectrum, emitting surfaces of very low work function are required. It has been found that a composite surface of caesium, caesium oxide, and silver has the lowest work function yet discovered, and this surface is accordingly used for "red-sensitive" photoelectric tubes, which must emit large numbers of electrons under the influence of red and infrared light.

When light is allowed to strike a photoelectrically sensitive surface, it is found that the maximum energy of the emerging electrons (which can be determined by measuring the potential difference necessary to arrest their motion) is proportional to the frequency of the light. This is a logical consequence of the supposition that the emerging electrons receive their energy from the light quanta, whose energy is also proportional to the frequency of the light. It is also found, in agreement with the theory that, when the *intensity* of the light is increased, the energy of the electrons is not increased, but the number of electrons emitted is increased in direct proportion to the increase in inten-

sity. The total photoelectric emission thus depends directly on the strength of the incident light.

If the strength of the light is kept constant, but its color varied, then the changes in electron emission are less easy to predict. Theoretically the number of electrons emitted per unit of light energy should decrease as the color is changed in the direction of the ultraviolet (Fig. 18), and to a limited extent this is observed. In practical cases, however, the changes in emission with changes in color can be expressed only in terms of an empirical (experi-

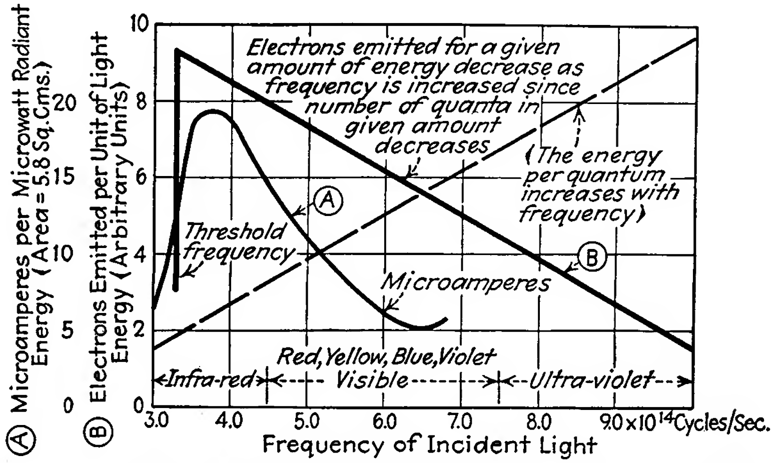


FIG. 18.—Actual (*A*) and theoretical (*B*) photoelectric emission *vs.* frequency of incident light. Curve *A* is that of a caesium oxide surface.

mentally determined) *spectral response curve*, which gives the number of amperes emission per square centimeter, when there is one lumen of incident light, for all the spectrum colors. Typical examples of these spectral response curves are given in Fig. 18 and in Chap. VIII, which describes practical photosensitive tubes.

An important descriptive term applied to photoelectric surfaces is the *threshold frequency*, which is the lowest frequency of light which will excite emission from the surface. As the frequency of the light is lowered toward the red end of the spectrum, the energy of the light quanta decreases in proportion, and at the threshold frequency the energy present in each quantum falls below the required work-function energy necessary to remove electrons from the surface. Table III gives the threshold fre-

quencies for various photoelectric surfaces, together with the corresponding work functions; it will be noted that the lower the work function, the lower the threshold frequency. The frequency of the various forms of visible and invisible light, together with the corresponding colors and wave lengths, is given in Fig. 19. The threshold frequencies of several metals and composite surfaces are included.

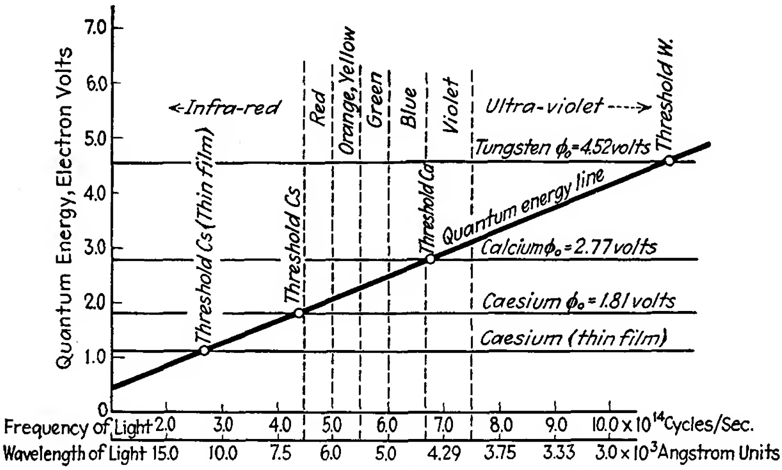


FIG. 19.—Relation between work function and threshold frequency. The intersection of the quantum-energy line with the work-function value determines the threshold frequency and wave length.

The relationship between threshold frequency and the work function of the metal can be shown in the following manner: The work-function energy (expressed in volts) is ϕ_0. The charge on the electron, in e.s.u., is e. The energy W_r required to remove the electron is then (from Eq. 7a, page 35):

$$W_r = \frac{e\phi_0}{300} \text{ ergs.} \tag{7b}$$

This energy must be supplied by the light quantum. The quantum energy W_q, in ergs, is

$$W_q = hf \text{ ergs,} \tag{7c}$$

where f is the frequency in c.p.s. and h is Planck's constant, 6.54×10^{-27} erg-sec. When emission occurs, the quantum

energy must be equal to, or greater than, the work-function energy, *i.e.*,

$$hf \geqq \frac{e\phi_0}{300} \tag{7d}$$

TABLE III.—CONSTANTS OF PHOTOELECTRIC EMITTING SUBSTANCES

Substance	Work function, volts	Threshold wave length, angstrom units[1]	Atomic radius, cm.[2]
Caesium (Cs) (thin film)	1.81	10,000	2.37×10^{-8}
Magnesium (Mg)	2.42	3,430	
Nickel (Ni)	5.01	3,050	1.25×10^{-8}
Platinum (Pt)	6.26	2,570	1.39×10^{-8}
Potassium (K)	2.24	7,000	
Sodium (Na)	2.46	5,830	
Tungsten (W)	4.52	2,300	1.36×10^{-8}

or, rearranging,

$$f_0 = \frac{e\phi_0}{300h} \tag{12}$$

This frequency f_0 is the threshold frequency, the lowest light frequency which will excite photoelectric emission from a surface whose work function is ϕ_0 volts.

Problem 3. An incandescent lamp is fitted with an infrared filter, so that the light supplied by it contains frequencies not higher than 4.5×10^{14} c.p.s. What work-function energy corresponds to this frequency? Referring to Table III, determine what substances are photoelectrically sensitive to this light.

Given:

$f_0 = 4.5 \times 10^{14}$ c.p.s.
$e = 4.8 \times 10^{-10}$ e.s.u.
$h = 6.54 \times 10^{-27}$ erg-sec.

To find:

$$\phi_0 = \frac{300h \times f_0}{e} \qquad \text{[from Equation (12)]}$$

$$= \frac{300 \times 6.54 \times 10^{-27} \times 4.5 \times 10^{14}}{4.8 \times 10^{-10}} \text{ volts} = \mathbf{1.85 \text{ volts.}}$$

[1] KOLLER, "The Physics of Electron Tubes," 2d ed., p. 168, McGraw-Hill Book Company, Inc., New York, (1937).
[2] DUSHMAN, *Rev. Mod. Phy,* **2**, 394 (1930).

Caesium and composite caesium surfaces are suitable. The caesium type of phototube was developed primarily for use with incandescent-light sources whose light output is rich in infrared light and deficient in the violet.

11. Secondary Emission.—Almost any surface can be made to emit electrons when bombarded by other electrons of sufficiently high kinetic energy (Fig. 20). Bombardment by positive charges, such as positive ions or protons, will likewise produce emission, provided only that the energy transferred from the bombarding particles is greater than the work-function energy of the surface. For this reason, surfaces with low-work-functions energy are especially suitable for obtaining secondary emission. In practical tubes making use of secondary emission, (see page 214), the surface used is usually caesium oxide on silver.

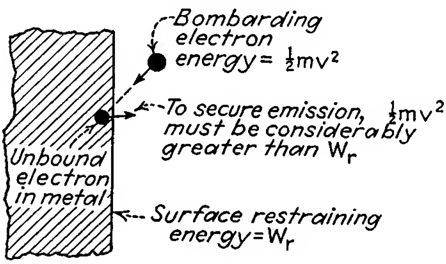


FIG. 20.—Secondary emission.

In most cases, however, the presence of secondary emission is to be avoided, since it is a type of emission which subtracts from the useful electron flow through the tube. The use of metals having high work functions, and special surface treatments which insure low operating temperatures, are both effective in reducing secondary emission.

A highly interesting aspect of secondary emission is the fact that a high-speed bombarding electron may liberate from the surface as many as ten "secondary" electrons. This amounts to a multiplication of the electron flow by a ratio as great as 10 and is put to use in some of the newer experimental tubes which are used as current-multiplier devices.

12. High-field Emission.—Whenever there are positive charges in the vicinity of a metal surface, but not actually on the surface, these charges exert an attracting force on the unbound electrons within the surface as shown in Fig. 21. If the concentration of positive charge is great enough or near enough the surface, it will

succeed in counterbalancing the restraining forces of the metal surface, and the electrons will thereupon be emitted from the surface. Under vacuum conditions the necessary concentration of positive charges may be set up by connecting a very high voltage difference (usually many thousands of volts) between two electrodes, the negative one of which will then emit electrodes under the attracting influence of the other. When gas is present near the emitting surface, the positive-charge concentration may arise from the production of positive ions in the gas itself. This effect is explained in detail in Chap. V, in which electron currents in gases are discussed.

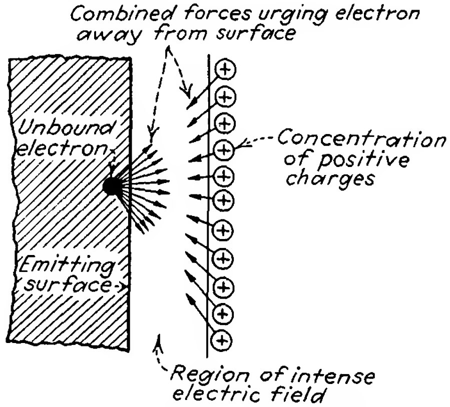


FIG. 21.—Field ("cold") emission.

The principal use of high-field emission in practical electron tubes is in the mercury arc. A pool of mercury which acts as the negative electrode of a tube will provide the emission of electrons through a "cathode spot," a rapidly moving and highly luminescent spot on its surface. Since the measured temperature of the cathode spot is much lower than would be required to produce the observed emission by thermionic means, it is believed that the emission results from the electric field produced by a high concentration of positive mercury ions near the surface of the pool. The rapid motion of the spot is due to the bombardment of the positive ions, which literally "blow" the spot about, but the energy of bombardment is not high enough to produce appreciable secondary emission.

13. Measurement of Work Function from Emission Data.—Since the work function of the emitting surface is so intimately

connected with the ability of the surface to produce free electrons, it is to be expected that the work function of different surfaces might be measured from the emission produced by various means. Many work functions have been measured by noting with precision the photoelectric threshold frequency of the surface in question. Likewise data giving the amount of emission obtained thermionically at various temperatures have also revealed work-function values. To illustrate the method in the latter case, consider the following problem:

Problem 4. A tungsten wire of unknown composition emits 0.1 amp./sq. cm. at a temperature of 1900°K. Find the work function of the tungsten surface and determine whether the tungsten is pure, or contaminated with a substance of lower work function.

Given:

$A = 60.2. \qquad I_s = 0.1$ amp./sq. cm.

$T = 1900$°K.

$\epsilon = 2.718.$

$k = 0.863 \times 10^{-4}$ volt per degree.

To find ϕ_0 in the expression:

$$I_s = AT^2\epsilon^{\frac{-\phi_0}{kT}}$$

$$0.1 = 60.2 \times (1900)^2 \times 2.718^{\frac{-\phi_0}{0.863\times 10^{-4}\times 1900}}$$

$$2.718^{\frac{-\phi_0}{0.16}} = \frac{0.1}{60.2 \times (1900)^2} = 4.6 \times 10^{-10}.$$

$$\frac{\phi_0}{0.16} \times \log \frac{1}{2.718} = \log 4.6 + 10 \log \frac{1}{10}.$$

$$\frac{\phi_0}{0.16} = \frac{0.663 - 10}{-0.434} = 21.5.$$

$$\phi_0 = 3.44 \text{ volts.}$$

Since the work function of a pure tungsten surface is 4.52 volts, the sample must be contaminated. Thoriated tungsten has work functions from 2.63 to 4.52, depending on the percentage of the surface covered with metallic thorium. The chances are, therefore, that the sample is thoriated tungsten.

Problems

1. Find the voltage equivalent to an electron energy of 6.0×10^{-12} erg. Find the electron-energy equivalent to a voltage of 3.2 volts.

2. A tungsten surface has a work function of 4.5 volts. Find the energy required to release an electron from the surface. Find the approximate diameter of the tungsten atom.

3. A barium-strontium oxide surface has a work-function voltage of 1.5 volts. Find the energy required to remove an electron from the surface.

Find the velocity, in the direction of the surface at which the electron must travel to free itself.

4. Find the total emission in amperes from a thoriated tungsten wire (100 per cent thorium layer) operated at 1900°K., whose dimensions are 10 cm. long and 0.01 cm. in diameter.

5. What area of barium-strontium oxide would be required to provide a total emission of 100 amp. if operated at 1200°K.?

6. Calculate the threshold frequencies for sodium, potassium, and platinum from the given values of work-function voltage.

7. A caesium-oxide-silver surface ($\phi_0 = 1.8$ volts) bombarded by electrons of 100 volts' energy releases an average of 5 secondary electrons for every bombarding electron. What is the efficiency of electron-energy conversion? If the emitted secondary electrons are caused to bombard another surface after gaining 100 volts' energy each, how many secondaries arise after the second bombardment? How many would arise after 10 successive bombardments. (This cumulative electron multiplication is utilized in multiplier tubes—see page 214.)

Bibliography

KOLLER: "Physics of Electron Tubes," McGraw-Hill Book Company, Inc., New York, Chaps. I to V, XII.

REIMANN: "Thermionic Emission," John Wiley & Sons Inc., New York, 1934.

HUGHES and DUBRIDGE: "Photoelectric Phenomena," McGraw-Hill Book Company, Inc., New York, 1932.

DUSHMAN: "Thermionic Emission," *Rev. Mod. Phys.*, **2**, 381 (1930).

ALLEN: "Photoelectricity," Longmans, Green & Company, London, 1925.

CHAPTER IV

THE CONTROL OF FREE ELECTRONS IN A VACUUM

14. The Electric Field and Its Properties.—In Chap. II (page 25) it is pointed out that the principal means of controlling the motion of free electrons is the repulsion force exerted by a concentration of other electrons on an electrode located in the space near the free electrons. In undertaking a more detailed study of this method, it is convenient to describe the repulsion force in terms of the *electric field* which exists in the space surrounding any electric charge.

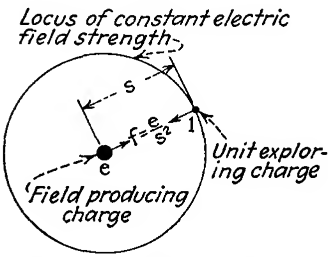


FIG. 22.—Electric field surrounding a single electron of charge e.

Consider an electric charge of e e.s.u. situated at a definite point in space. Around this charge exists a field of force, since any other charge placed at any other point in space will experience a force due to the presence of this first charge. The *strength of the field* is equal to the force in dynes experienced by a *unit* charge (1 e.s.u.). Suppose the field-producing charge e and the unit "exploring" charge are separated by a distance of s cm. (Fig. 22), then according to Equation (5) (page 24), the force exerted on the unit charge is

$$f = \frac{e_1 e_2}{s^2} = \frac{e \times 1}{s^2} \text{ dynes.}$$

This force is the field strength at a point s cm. from the charge e. The field possesses radial symmetry, since s can be measured away from e in any direction; but the strength of the field decreases with the square of the distance s; hence the electric field is nonuniform. The field strength is usually given the symbol X; hence for a point charge e the field is described by

$$X = \frac{e}{s^2}.$$

This type of field is not used in practical electron tubes to any extent. Rather, a uniform field is used as the basis for most designs. As shown in Chap. II (page 24), a uniform field is produced between two large parallel plates when a battery is connected between them, so that one plate is covered with a uniform distribution of electrons, leaving a uniform deficiency of electrons on the other (Fig. 23). The force on a unit charge

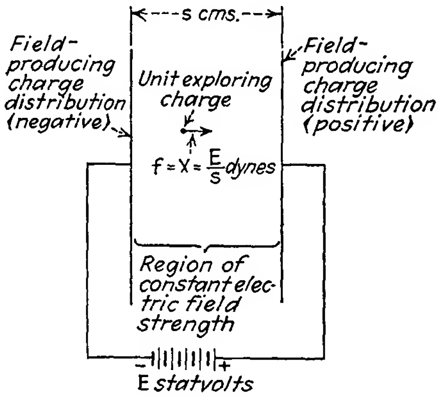


FIG. 23.—Electric field between two large parallel plates.

(field strength) between these plates is the same at any point. It is given by

$$X = \frac{E}{s}, \tag{13}$$

where X is the force in dynes exerted on a unit charge by a potential difference of E statvolts applied between large parallel plates separated by s cm. The force in this case is thus numerically equal to the number of statvolts voltage divided by the number of centimeters between the plates.

Even in nonuniform fields, the force at any point on a unit charge (the field strength X) can be expressed as the number of statvolts voltage change per centimeter, except that in the nonuniform field the change in voltage must be taken over a very small change in distance, since the force changes rapidly as the distance changes. If the field is nonuniform, the field strength is

$$X = \frac{dE}{ds}, \tag{14}$$

where dE stands for a very small voltage change and ds is the corresponding very small distance change. In practice it is convenient to make a graphical plot of the nonuniform change in voltage against the change in distance. The slope of this plotted curve, at any point, is then proportional to the field strength at that point as shown in Fig. 24.

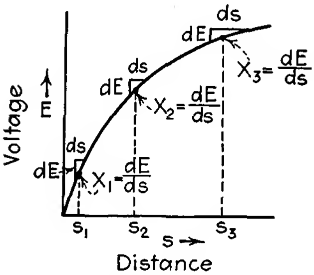


FIG. 24.—When the voltage distribution curve is known, the electric field strength at any point on the curve may be determined by taking the slope of the curve at that point. In order of decreasing field strength, the points above are $s_1(X_1)$, $s_2(X_2)$, $s_3(X_3)$.

Since the field strength is thus a voltage divided by a distance, it follows that a voltage is a force multiplied by a distance, *i.e.*, in the uniform-field case

$$E = X \times s \tag{15}$$

and in the nonuniform case

$$dE = X \times ds \tag{16}$$

The dimensions of voltage, force times distance, are thus the same as those of energy.

15. The Effect of an Electric Field on a Single Electron.—The effect of a uniform field on a single electron has been examined in Chap. II (pages 23 to 26). The electron moves with constant acceleration (constant increase in velocity) from the negative to the positive plate, acquiring at the latter plate a final velocity which can be calculated. This velocity, with the mass of the electron m determines the final kinetic energy. If the voltage applied is E statvolts, and the distance between the plates is s, then the force acting on an electron of charge e is

$$f = \frac{eE}{s} \text{ dynes.} \tag{6a}$$

The energy at the end of the electron flight is

$$f \times s = eE. \tag{7}$$

This energy appears as kinetic energy; hence,

$$\text{K.E.} = \tfrac{1}{2}mv^2 = eE \text{ ergs} \tag{17}$$

and the final velocity at the end of the flight v is

$$v = \sqrt{\frac{2eE}{m}} \text{ cm./sec.} \tag{18}$$

Now the electron current carried by a single electron is directly proportional to its average velocity, $v_{av.}$ The velocity given in Equation (18) is the final velocity, but since the electron starts from rest and increases in velocity at a constant rate, the average velocity is just one-half the final velocity. Hence the current flow between two parallel plates (see Fig. 25), carried by a single electron, is the following:

$$i = kv_{av.} = \frac{kv}{2} = k\sqrt{\frac{eE}{2m}} \tag{19}$$

The current thus increases proportionately to the square root of the applied voltage difference.

When more than one electron participates in the current flow, then the mutual repulsion of the electrons on each other comes into play and the expression in Equation (19) does not hold. This latter case is treated in Sec. 17 (page 57).

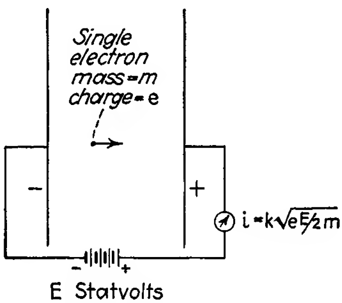


FIG. 25.—The current carried by a single electron between parallel plates depends on the square root of the applied voltage.

16. The Millikan Oil-drop Experiments.—As an illustration of the action of an electric field on single electric charges, we consider Millikan's oil-drop experiment, by which the charge on the electron was determined. In this experiment, the charges were not in a vacuum, so that the discussion of Sec. 15 does not hold in its entirety, but the principles are brought out nevertheless.

Two metal plates are placed one above the other, parallel and separated by about 0.5 cm. A battery is connected to these plates so that the bottom plate is negative. The voltage applied is controlled by a voltage divider, as shown in the diagram (Fig. 26). The top plate has several very fine holes in it. Above this plate, oil is sprayed from an atomizer. The oil drops, irradiated with X rays, so that they are charged electrically, fall through

the fine holes into the electric field between the two plates. Here each oil drop finds itself acted on by two forces, the force of gravity urging it downward, and the force of the electric field, urging it up (assuming the charge on the droplet to be negative).

In the absence of the electric field (when the battery is disconnected) the drops fall through the air between the plates, and since the air is a viscous medium, they fall with constant velocity. When the electric field is applied, urging the charged droplets upward, the drops move against the force of gravity, provided the applied electric field is large enough. In both cases (upward and downward motion) the velocity of the droplet is proportional to the force producing it. When gravity acts alone, the force is

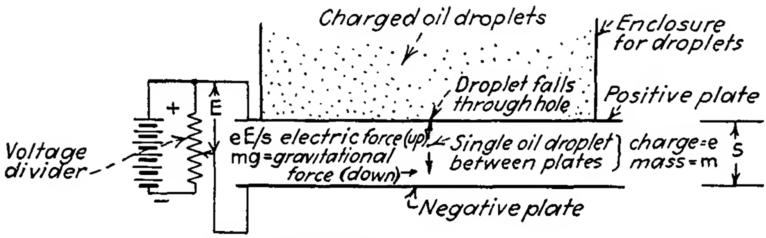


FIG. 26.—The Millikan oil-drop experiment.

mg where m is the mass of the oil drop and g is the gravitational constant. When the field is working against gravity, the net force is $\frac{eE}{s} - mg$, where e is the charge on the droplet, E the applied voltage in statvolts, and s the separation between the plates. Since these forces are proportional to the velocities they produce, the following equation holds:

$$\frac{(eE/s) - mg}{mg} = \frac{v_{fg}}{v_g} \tag{20}$$

where v_{fg} is the net velocity produced by both electric field and gravity and v_g is the velocity produced by the gravity alone. All the quantities in Equation (20) are measurable except the charge e. Hence e, the charge on the oil drop, can be calculated in terms of the measurable quantities.

The velocities v_g and v_{fg} are measured by watching the droplets through a microscope under a strong crossbeam of light, and timing their rates of fall. E and s are both known, and g is a

known constant. The mass of the droplet is not so easily determined, however. The density of the oil is known, so that if its volume ($\frac{4}{3}\pi r^3$, where r is the radius) is known, the mass can be calculated. The radius of the oil drop is determined by the application of Stokes' law, which says that the force experienced by a sphere falling through a viscous medium is $6\pi\mu r v_g$, where μ is the coefficient of viscosity of the medium (air in this case), r is the radius of the drop, and v_g the velocity of fall. The radius of the droplet is computed, by equating the Stokes' force to the force of gravity.

This experiment was repeated thousands of times by Millikan and his assistants,[1] and the charge on each drop calculated and recorded. It was found that the charge was always some whole multiple of (equal to, twice as great as, three times as great as, etc.) the elementary charge, and that this elementary charge had a value of $4.770 \pm 0.005 \times 10^{-10}$ e.s.u. The experiment thus determined the value of the charge on the electron and showed that this charge was the same for all electrons, since the observed oil-droplet charges were always whole multiples of the basic charge. In 1935 it was discovered that the value of μ used by Millikan was in error, and the new value of the electronic charge, computed from X-ray measurements, and recomputed from Millikan's data with the new value of μ, is $4.803 \pm 0.003 \times 10^{-10}$ e.s.u.

17. The Action of an Electric Field on a Group of Electrons; Space Charge.—When more than one electron is situated between parallel plates, it is necessary not only to consider the force produced by the charge on the plates but also the forces produced by each electron in the space on the other electrons in the space. The mathematical treatment of this case is complicated, but it can be handled by the methods of calculus.

The basis of the analysis is Poisson's equation, which states that the rate at which the force on the electron changes at any point is proportional to the concentration of electrons at that point (refer to Fig. 27). The concentration of electrons is given by ρ, and is expressed in e.s.u. per cubic centimeter. The velocity of each electron v, as computed from Equation (18), is $v = \sqrt{2eE/m}$. Now the concentration of electrons times their

[1] For an account of this work see R. A. Millikan, "The Electron," University of Chicago Press, Chicago, 1924.

velocity is the current flow per unit area, *i.e.*, $I = \rho v = \rho\sqrt{2eE/m}$. Therefore the concentration of electrons is $\rho = I/\sqrt{2eE/m}$. Now the rate at which the force in the electron changes is written, in the symbols of calculus, as d^2E/ds^2, and this is equal to 4π times the electron concentration. That is,

$$\frac{d^2E}{ds^2} = \frac{4\pi I}{\sqrt{2eE/m}}. \tag{21}$$

Equation (21) is a total differential equation whose solution is an equation involving E, the voltage between the plates and I, the

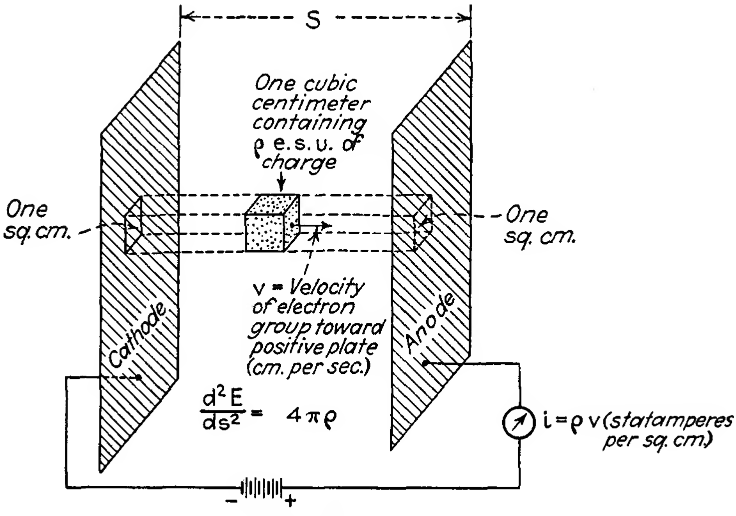


FIG. 27.—Relations determining the current carried by a large group of electrons passing between parallel plates.

resulting current flow. When the solution is substituted in Equation (21) by the rules of calculus, an identity results. The correct solution is

$$I = \frac{\sqrt{2e/m}E^{3/2}}{9\pi s^2}. \tag{22}$$

The substitution of Equation (22) in Equation (21) will reveal[1] that it is the correct solution.

[1] The substitution is as follows:

Equation (22) states that the current produced depends directly on the $\frac{3}{2}$ power of the applied voltage and decreases with the square of the distance separating the plates. When all the constants are evaluated, Equation (22) becomes[1]

$$I = \frac{2.34 \times 10^{-6}\, E^{3/2}}{s^2} \text{ amp./sq. cm.} \qquad (23)$$

This equation gives the current I, in amperes per square centimeter of plate surface, flowing between two large parallel plates separated s cm. and with a voltage of E volts (not statvolts) applied between them (Fig. 28). When the voltage is doubled, the current increases $2^{3/2} = 2.8$ times; when the voltage is tripled, the current increases $3^{3/2} = 5.2$ times, and so on. On the other hand, if the separation between the plates is doubled, the current becomes one-quarter of its first value, and so on. The equation shows, therefore, that the current flow (electron motion) between parallel plates can be controlled either by varying the applied voltage, or the separation between the plates, or both.

The above derivation of the $\frac{3}{2}$-power law assumes that an infinite supply of electrons is available, *i.e.*, that the negative plate is emitting electrons without limit. In practice the negative plate must emit more current per square centimeter than is required by Equation (23), otherwise the equation does not hold.

$$I = \frac{\sqrt{2e/m}}{9\pi s^2} E^{3/2}. \qquad (22)$$

$$E = \left(\frac{9\pi I s^2}{\sqrt{2e/m}}\right)^{2/3}. \qquad (22a)$$

$$\frac{d^2E}{ds^2} = \left(\frac{9\pi I}{\sqrt{2e/m}}\right)^{2/3} \frac{4}{9s^{2/3}} = \left(\frac{8\pi I}{3\sqrt{2e/m}\, s}\right)^{2/3}. \qquad (21a)$$

But, according to Equation (21),

$$\frac{d^2E}{ds^2} = \frac{4\pi I}{\sqrt{2e/m}} \frac{1}{\sqrt{E}}$$

$$= \frac{4\pi I}{\sqrt{2e/m}} \times \left(\frac{\sqrt{2e/m}}{9\pi I s^2}\right)^{1/3} \text{ (from 22}a\text{)}$$

$$= \left[\frac{64\pi^2 I^2}{9(2e/m)s^2}\right]^{1/3} = \left[\frac{8\pi I}{3\sqrt{2e/m}\, s}\right]^{2/3} \qquad (21b)$$

Since (21a) and (21b) are the same, the identity is established.

[1] This equation was originally derived by Child, and is sometimes called Child's law. See *Phys. Rev.*, **32**, 498 (1911).

If fewer electrons are being emitted, then the current is limited by the number emitted, and increasing the applied voltage or decreasing the separation between the plates will not increase the current.

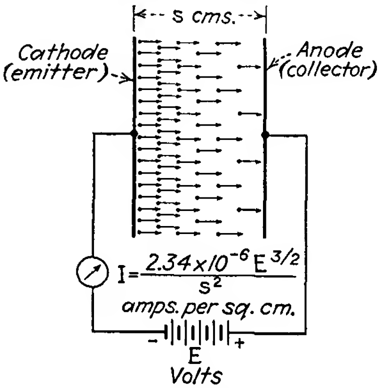


FIG. 28.—Current flow *vs.* applied voltage and plate separation, expressed in practical units.

The ³⁄₂-power Law for Cylindrical Electrodes.—A very practical arrangement of electrodes used in electron tubes consists of two coaxial cylinders, shown in Fig. 29 the inner of which is usually the negative electrode, the outer the positive. In this case the ³⁄₂-power law takes the form[1]

$$I = \frac{14.68 \times 10^{-6} E^{3/2}}{r\beta^2} \text{ amp./cm. length,} \tag{24}$$

where r is the radius of the outer cylinder, and β^2 is a quantity

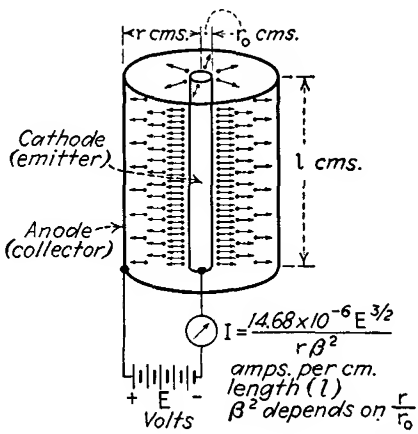


FIG. 29.—Current flow between concentric cylinders, in terms of applied voltage, anode radius, and β^2 (see Fig. 30).

which depends on the ratio of the diameter of the outer cylinder to that of the inner cylinder. Tables of β^2 are available (see

[1] The cylindrical form is due to Langmuir. Langmuir also showed that the current between electrodes of any shape and separation varies as the ³⁄₂-power of the applied voltage. See *Phys. Rev.*, **2**, 450 (1913); **22**, 347 (1923).

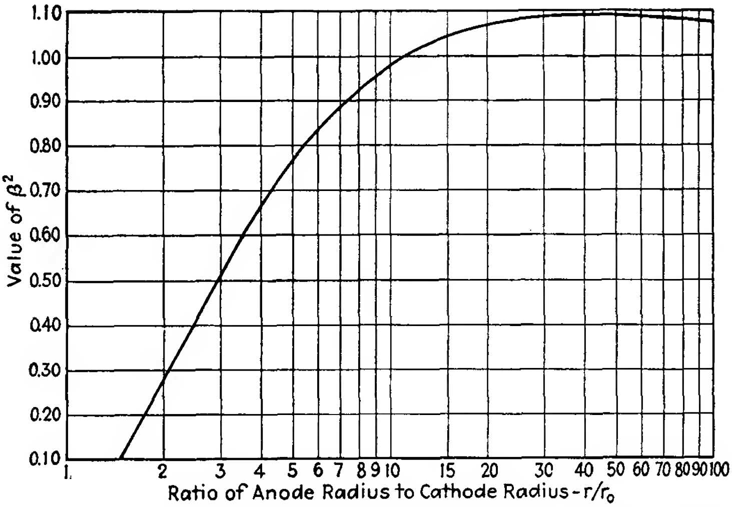


FIG. 30.—Values of β^2 in terms of ratio of anode radius, r, to cathode radius, r_0.

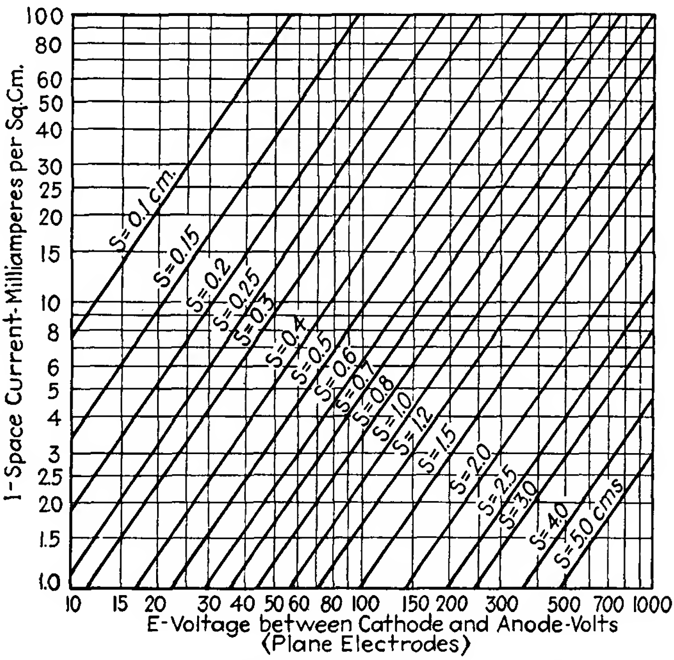


FIG. 31.—Space current *vs.* applied voltage for plane parallel electrodes. [Calculated from Eq. (23).]

Fig. 30), but for practical cases it is usually sufficient to take the value $\beta^2 = 1$, which holds within a few per cent error when the outer cylinder has a diameter ten times as great as the inner. Note that the current in this case depends on the first power of the radius, and that the current is expressed in amperes *per centimeter length* of cylinder.

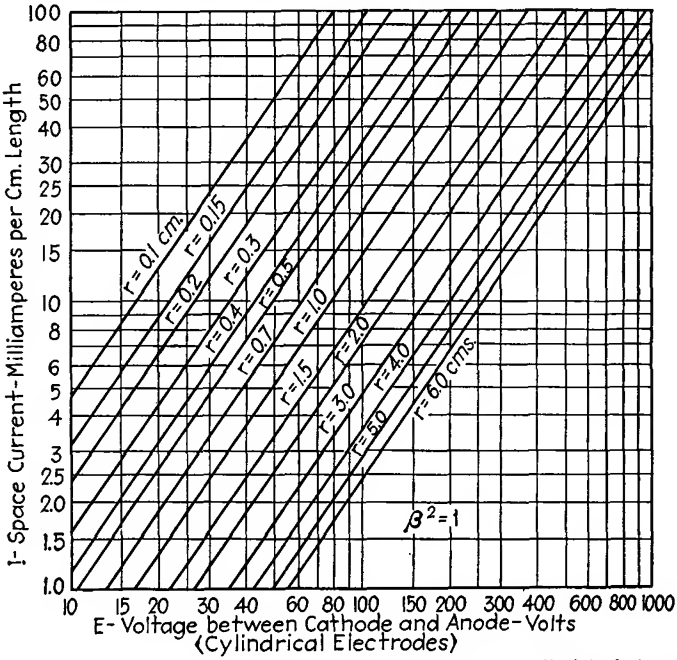


FIG. 32.—Space current *vs.* applied voltage for concentric cylindrical electrodes. [From Eq. (24).]

In both cases (plane and cylindrical electrodes), the current I enters the positive electrode and is conducted to the external circuit. The current flowing in the circuit thus depends, first, on the electron-emitting ability of the negative electrode. If this electrode has a low work function and if the releasing energy (thermal, photoelectric, field, or secondary) is great enough, then the supply of electrons will be plentiful. The current flowing then depends on the voltage applied between the two electrodes, and on the geometry of the electrodes, in particular on the separation between them. The current flow can be computed in two important practical cases by the use of Equations

(23) and (24), for plane and cylindrical electrodes, respectively. Representative values of I have been computed in Figs. 31 and 32.

18. Grid Control of Electron Currents in a Vacuum. Potential Distribution in the Interelectrode Space.—The foregoing means of controlling electron-current flow (by emission characteristics, by applied voltage, and by changing the geometry of the electrodes) are limited in their ability to control the current over wide ranges. A much more effective method consists in applying a voltage in the space midway between the electrodes, by inserting a *grid* electrode through which the electrons can pass, and by connecting a battery between this grid and the negative electrode (Fig. 33). Relatively small changes in this battery voltage will produce very large changes in the current flowing between the negative electrode and the outer positive electrode. This is the action on which electronic amplification (discussed in Chap. I, page 6) depends, and it is therefore a very important part of the study of electron action.

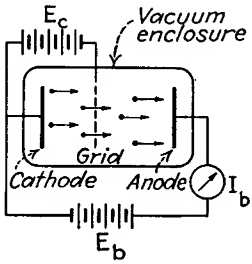


FIG. 33.—Grid control of space current.

To aid in understanding the action of the grid in controlling electron currents, it is useful to use a plot of the *voltage distribution* between the electrodes. Such plots are shown in Figs. 34 to 36, for the case of two parallel plane electrodes connected to a battery. If only one electron (or at most a small number of them) is present between the plates, then the force on each electron is the same at every point between the plates, and the voltage distributes itself uniformly over the space between the plates, as shown by the straight line in Fig. 34*a*. The force on the electron is the slope of this line which coincides with the line itself; hence the force is the same whether the electron is near the first or the second plate, or anywhere in between.

When a large group of electrons is present in the interelectrode space, however, the forces between the electrons come into play. The first electrons emitted are urged forward by the concentration of electrons in the space behind them, while those behind are urged backward by those ahead. Hence the electrons nearest the positive plate are acted on by a greater total force and the slope of the voltage curve is greater near the positive plate than

it is near the negative plate, as shown in Fig. 34*b*. This is the voltage distribution encountered in most practical vacuum tubes, and the one from which the current in Equation (23) is calculated. Since the total force on electrons near the positive

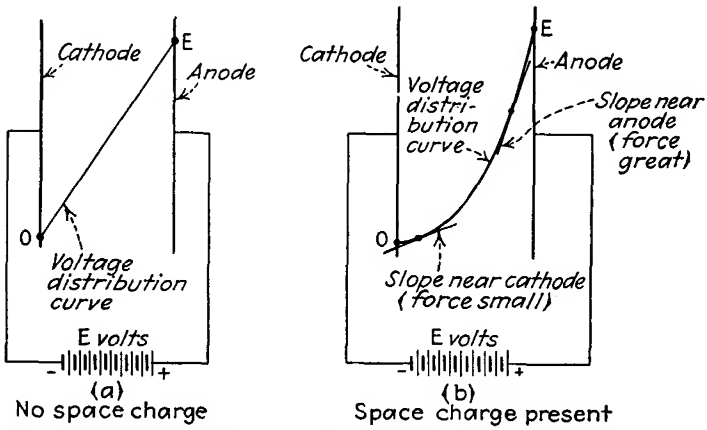


FIG. 34.—Voltage-distribution curves between plane electrodes: (*a*) when few electrons are present; (*b*) when large numbers of electrons are present.

plate is greater than that on the electrons near the negative plate, the former move faster than the latter, and the electron concentration or density thins out as they near the positive plate.

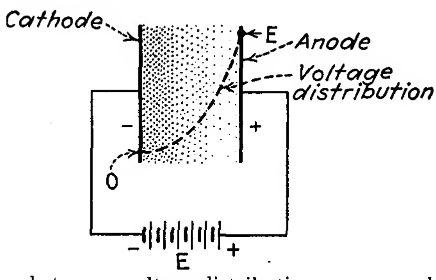


FIG. 35.—Relation between voltage-distribution curve and electron-density distribution.

The voltage distribution thus gives rise to the distribution of electrons shown in Fig. 35.

If a grid is situated between the negative electrode (cathode) and the positive electrode (anode) and if this grid is connected to a battery of E_c volts while the anode is connected to a battery of E_b volts, then the voltage distribution between the anode and

cathode depends not only on the anode voltage but on the grid voltage as well. If the grid is made positive with respect to the cathode by a sufficient amount, then the voltage-distribution curve is "lifted" by the grid, as shown in Fig. 36*a*. If the grid is made negative with respect to the cathode, then the voltage-distribution curve is depressed, as in Fig. 36*b*. This lifting or depressing of the curve changes the slope of the voltage-distribution curve at the cathode surface, and hence changes the force acting on the electrons (the force being proportional to the slope). When the grid is positive, the force is great; but when the grid is

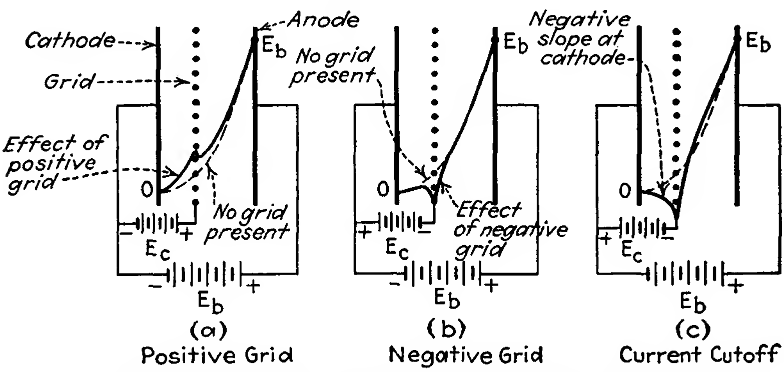


FIG. 36.—Effect of grid voltage in raising and depressing the voltage-distribution curve.

negative, the force is decreased; and it may be decreased to zero if the voltage on the grid is negative enough. If this happens (Fig. 36*c*), the electrons are not urged away from the emitting cathode at all, and the current flow ceases entirely. This is the "cutoff" point frequently referred to in discussing vacuum-tube characteristics and operation. If the grid is negative, moreover, it cannot collect electrons, and hence it can control the electrons without interfering with the current flow itself. This makes grid control a highly efficient and sensitive method of changing electron-current flow. Consequently in most cases the grid is operated at a potential whose average value is negative.

19. The Formation of Electron Beams and Their Control by Electrostatic and Magnetic Forces.—The electron currents considered thus far are those flowing over areas situated between charged plates. It is possible to produce quite a different form of electron flow by suitably choosing the shape and arrangement

of the electrodes. If the cathode (electron emitter) frees electrons which are drawn through a small hole, and if the anode which exerts the attractive force is cylindrical in shape, as shown in Fig. 37, then the electrons will flow from the hole in the form of a beam, to be collected eventually by the anode. Such a beam of high-speed electrons is called a *cathode ray* (after Crookes'

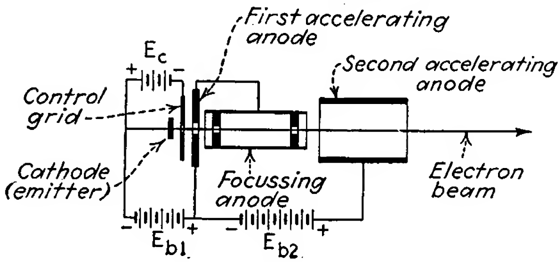


FIG. 37.—"Electron gun" structure, used for forming a narrow beam of electrons.

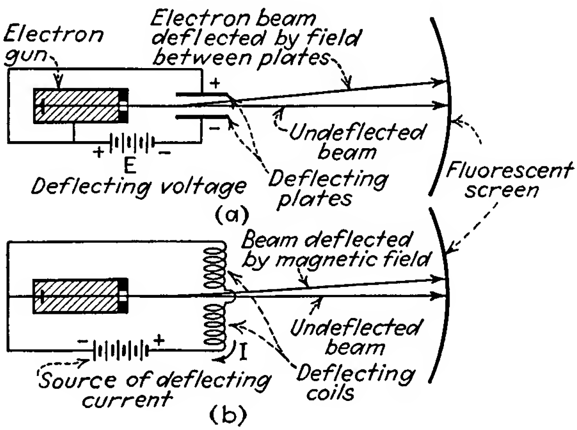


FIG. 38.—Methods of deflecting electron beams: (*a*) electric and (*b*) magnetic.

discovery, page 18). If the beam of electrons hits a part of the glass container which is coated with a fluorescent material, such as artificial Willemite (zinc orthosilicate), the screen will glow with a bright spot where the electron beam hits. Any motion of this fluorescent spot is an indication of a change in the direction of the beam and hence is a very convenient indication for the study of the control of electron beams.

Two principal methods of controlling the direction of an electron beam are used, as shown in Fig. 38, One uses electric forces

produced by connecting a battery between two plates, one on either side of the beam. The other makes use of magnetic forces produced by sending current through two coils, one on either side of the beam.

Electrostatic Deflection.—If the electron beam passes between two parallel plates whose separation is s cm., then the force acting to move each electron toward the positive plate is eE/s dynes, where E is the applied voltage in statvolts. Each electron

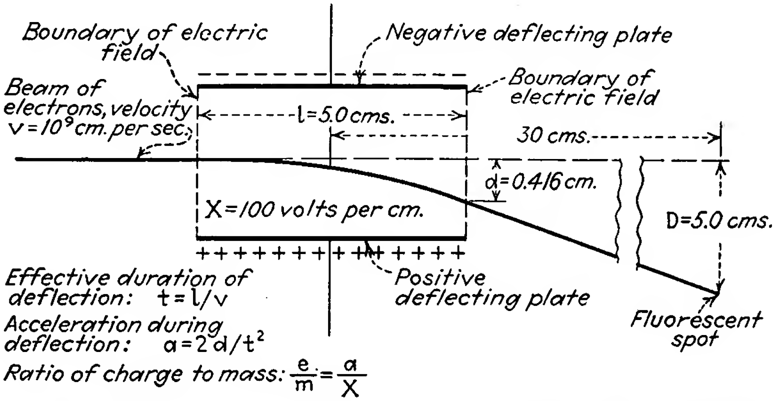


FIG. 39.—Deflection of an electron beam by an electric field. (Numerical values refer to Prob. 1, page 69).

remains within the influence of this force for $t = l/v$ sec. (equal to the length of the plates divided by the velocity of the electrons), and is deflected a distance $d = \frac{1}{2}at^2$, where a is the acceleration due to the force and is equal to the force eE/s divided by the mass m of the electron. This deflection is the distance moved at right angles to the beam. The change in direction taken by the beam is the angle ϕ in

$$\tan \phi = \frac{eEt^2}{msl} = \frac{eXt^2}{ml}$$

These relationships are shown in Fig. 39. From them can be calculated the change in motion of the spot in terms of the electron mass and charge, the applied voltage, the separation between the deflection plates and the distance to the fluorescent screen.

Magnetic Control of Electron Beams.—If two coils situated on either side of the electron beam (Fig. 38b) are fed such a current that the magnetic field between them is H electromagnetic units

(e.m.u.), then the magnitude of the force exerted on each electron in the beam is

$$f = H \times e \times v \text{ dynes,} \tag{25}$$

where e is the electron charge (in e.m.u.) and v is its velocity in centimeters per second. The direction of the force is at right angles both to the lines of magnetic force and to the direction of the electron motion. If the force acts only for a short time (*i.e.*, if the speed of the electrons along the beam is very great compared with the deflection speed they acquire at right angles to the beam), then the amount of deflection can be calculated by the same method as is used for electrostatic deflection, *i.e.*, by computing the time the electron is under the influence of the force, the distance it moves at right angle to the beam motion, and comparing this distance with the extent of the magnetic field between the coils.

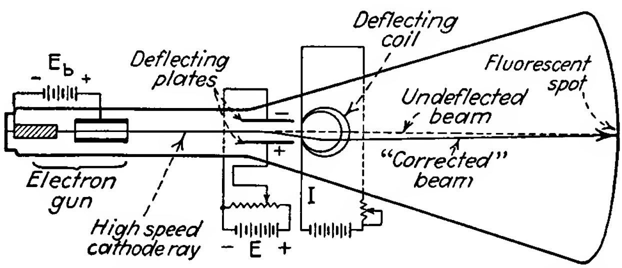


FIG. 40.—By using both electric and magnetic fields, the deflection caused by one field may be cancelled by the other. The velocity of the electrons in the beam may be computed from the values of the field strengths.

Conversely, if the amount of deflection be measured directly by the motion of the fluorescent spot, then any unknown quantity in the deflecting system may be calculated in terms of the other known quantities. In particular the values of the deflecting electric or magnetic fields may be calculated, and hence the voltages and currents producing these fields.

Crossed Fields. The Thomson Cathode-ray Experiment.—An interesting effect is obtained if both electric and magnetic deflecting forces are applied simultaneously in such a way that the magnetic deflecting force exactly opposes the electric deflecting force. This may be obtained by the arrangement of coils and plates shown in Fig. 40 and by adjusting the plate voltage and the

coil current until the beam does not suffer any deflection. The fields are then said to be crossed. Since the two forces are exactly opposed, they must equal each other in magnitude. Equating the electrostatic force from Equation (6*a*) and the magnetic force from Equation (25), we obtain

$$H \times e \times v = \frac{e \times E}{s}$$

and solving for the velocity v,

$$v = \frac{E}{Hs} = \frac{X}{H} \tag{26}$$

Since E, H, and s can be measured, the velocity can then be computed. Knowing this velocity, it is simple to calculate the deflection which this same beam would undergo under electric or magnetic fields alone, by the methods of the preceding paragraph. Or, if the deflection is measured, it is possible to compute the charge-to-mass ratio by the same process. This is the method employed by Sir J. J. Thomson[1] in determining the charge-to-mass ratio of electrons obtained from many different substances.

Problem 1. To illustrate the calculations involved in this determination of e/m, we may consider the following values. It is found that zero deflection of the beam occurs when the crossed fields are a magnetic field $H = 10$ e.m.u., and by an electric field X of 100 volts per centimeter, equal to 10^{10} e.m.u. units of electric field (1 ordinary volt is equal to 10^8 e.m.u. volts, or abvolts). The velocity of the electrons in the beam is then

$$v = \frac{X}{H} = \frac{10^{10}}{10} = 10^9 \text{ cm./sec.}$$

The magnetic field is now removed and the observed deflection at the cathode-ray screen is found to be 5.0 cm. (see Fig. 39). The distance from the center of the deflecting plates to the screen is 30 cm. The length of the deflecting plates is 5.0 cm. Find the acceleration produced by the electrostatic field on each electron in the beam.

Given:

$v = 10^9$ cm./sec.
$l = 5.0$ cm.
$D = 5.0$ cm.
$L = 30$ cm.

[1] *Phil. Mag.*, **44**, ser. 5, 293 (1897).

To find:

t (time during deflection period).

$$t = \frac{l}{v} = \frac{5.0}{10^9} = 5.0 \times 10^{-9} \text{ sec.}$$

d (distance moved during deflection period).

$$\frac{d}{\frac{1}{2}l} = \frac{D}{L} = \frac{5.0}{30.0}.$$

$$d = \left(\frac{5.0}{30.0}\right)2.5 = 0.417 \text{ cm.}$$

a (acceleration).

$$a = \frac{2d}{t^2}$$
$$= \frac{2(0.417)}{(5.0 \times 10^{-9})^2}$$
$$= \tfrac{1}{30} \times 10^{18} \text{ cm./sec./sec.}$$

From this calculated value of acceleration, we can calculate the value of e/m as follows:

$$X = {}^{100}\!/_{300} \text{ statvolt per cm.}$$
$$f = eX = ma.$$
$$\frac{e}{m} = \frac{a}{X}$$
$$= \frac{\frac{1}{30} \times 10^{18}}{{}^{100}\!/_{300}}$$
$$= \tfrac{3}{30} \times 10^{18} \text{ e.s.u./g.}$$
$$= 0.1 \times 10^{18} \text{ e.s.u./g.}$$

In this hypothetical example the value calculated is not, of course, the correct value, which is 0.53×10^{18} e.s.u./g.

Problems

1. Calculate the final velocity of a single electron which has accelerated from rest through a voltage difference of 300 volts. By what percentage will the final velocity increase if the voltage is increased to 600 volts?

2. Calculate the current flow between two parallel plates (the negative of which is an inexhaustible emitter of electrons), each 10 cm. on a side, separated by 5 cm. and connected to a battery of 100 volts. What is the current when the applied voltage is 2700 volts?

3. Two coaxial cylinders, the inner being an inexhaustible emitter of electrons, are connected to a 100-volt battery. A current flow of 1.0 ma. is required between them. Determine one set of possible dimensions which would insure this current flow (diameter and length of each cylinder).

4. Using the values found in Prob. 3, determine the change required in the applied voltage to produce a change of 0.1 ma. in the current flow.

5. How much deflection (distance measured at right angles to the beam) will a beam of electrons (traveling at 10^9 cm./sec.) suffer in passing between two parallel plates each 10 cm. long, separated by 4 cm. and connected to a 100-volt battery?

6. A beam of electrons moves (deflects) a distance of 0.5 cm. in passing a distance of 10 cm. through a magnetic field of 10 e.m.u. What is the velocity of the electrons in the beam, assuming the electron charge is 1.6×10^{-20} e.m.u. and its mass 0.9×10^{-27} g.? (HINT: The time of acceleration in the deflection direction is $t = l/v = \sqrt{2s/a} = \sqrt{2s/(Hev/m)}$.

7. The correct value of the ratio e/m is 0.53×10^{18} e.s.u. per gram. Using all the information given in Prob. 1 (page 69), determine the value of the magnetic field which would insure a balance with an electric field of 100 volts/cm.

CHAPTER V

ELECTRON CURRENTS IN GASES AND VAPORS

Introduction.—The control of electrons in a vacuum is accomplished under nearly ideal conditions. The electric field set up between two plates urges the electrons to move toward the positive plate, while each electron repels the near-by electrons. Only these two forces need be considered, and the resulting equation (the $3/2$-power law) expressing the motion of the electrons is correspondingly simple. But when electrons move through a gas-filled space, the situation is greatly complicated by the molecules of gas. These molecules may enter into collisions with the moving electrons, and they may exert repelling or attracting forces on them. In many practical cases there are several types of gas molecule present, and this adds still further to the number of events which may occur. The energy interchanges between the electrons and the gas molecules are complicated by the great difference in their masses, by the preference of the molecule for certain energy states, and by the fact that the energy may be mechanical, electrical, or in the form of light. The situation is so complex, in fact, that no complete mathematical analysis of electron currents in gases is possible. But a serviceable working theory, which accounts for most of the observed facts, is available.[1] The more important aspects of this working background are presented in this chapter.

20. The Neutral Atom, Its Structure and Energy Levels.—The experiments of Sir J. J. Thomson and the late Lord Rutherford, and the reasoning of Moseley and Bohr, have revealed the essential structure of the atom. Each gas atom consists of a dense core or nucleus whose net charge is positive, surrounded by a distribution of negative charge in the form of electrons. The normal gas atom has as much positive as negative charge; this is the "neutral" state in which atoms exist. Under special

[1] For a complete account of the theory of gas discharges, as of 1931, see K. T. Compton and I. Langmuir, *Rev. Mod. Phys.*, **2**, 123 (1930); and **3**, 191 (1931).

circumstances, however, the atom may assume more positive charge than normal; it is then called a positive ion and is capable of exerting electric force on near-by charges. Less commonly the electron can be made to assume more negative charge than normal, by the addition of one or more electrons to its outer boundary. The forces which negative and positive ions can exert on near-by charges are of the greatest importance in explaining the action of electron currents in gases and vapors.

The Bohr Atom. Electron Orbits and Energies.—There is no general agreement concerning the positions and motions of the electrons about the positive nucleus, but there is one traditional view, first stated by Nils Bohr in 1913,[1] which has the virtue of simplicity and a high degree of usefulness. For the purposes of the present discussion we can state that the atom acts as though the electrons were revolving around the nucleus in closed, nearly circular orbits, and that the radius of each electron orbit has a definite value. The atom is according to this view a miniature solar system, the nucleus corresponding to the sun, and the revolving electrons to the planets.

The "planetary" motion of the electrons produces important results. The motion endows the electron with a centrifugal force, with which it resists the attracting force of the positive nucleus, and the motion also endows the electron with kinetic energy. In addition to the kinetic energy, each electron possesses potential energy, owing to its position near the positive nucleus. The total energy of the electron, the sum of its kinetic and potential energies, is given by the equation

$$W = -\frac{eQ}{2r} \text{ ergs,} \tag{27}$$

where W is the energy in ergs possessed by an electron of charge e e.s.u. revolving in a circular orbit of radius r cm., when the positive charge on the nucleus is Q e.s.u. The energy thus depends on the radius of the orbit. As the radius is increased, the energy decreases; but since the energy in the orbit is negative, an increase in the orbit actually means an increase in the energy of the electron. In other words, if the electron energy is increased by supplying it, for example, with energy from a fast-moving external electron, the radius of the electron orbit must increase.

[1] BOHR, NILS, *Phil. Mag.*, **25**, 10 (1913); **30**, 581 (1915).

Bohr's basic contribution to the theory is the so-called quantum condition. According to this view, amply supported by experiments with the arc spectrum of hydrogen, the energy of the electron is restricted to certain definite values. Certain values of energy are "allowed"; all others simply do not exist. The value of the "allowed" energy states can be calculated from the equation

$$W_n = \frac{-2\pi^2 me^2Q^2}{n^2h^2} \text{ ergs,} \tag{28}$$

where W_n is the total energy of an electron of mass m g. and charge e e.s.u., revolving in a circular orbit around a nucleus of positive charge Q e.s.u., and h is Planck's constant (6.54×10^{-27} erg-sec.). The quantity n is a number which may have any integral value (*i.e.*, 1, 2, 3, etc.), but not any fractional value. Since m, e, h, and E are known, the various possible values of energy may be computed from Equation (28) and the corresponding orbit radii from Equation (27).

Since the energy and the radius are related by Equation (27), it follows that if only certain values of energy are "allowed," then only certain radii are possible. If the radius of the orbit changes, owing to any cause, the electron must travel from one allowed radius to another allowed radius, *i.e.*, the electron must be conceived as making sudden jumps between the possible orbits. Each orbit jump is accompanied by the addition or subtraction of a definite amount of energy to or from the electron. This rather strange behavior is now well established in atomic theory.

As a corollary to the above behavior it follows that the internal energy of an atom may be increased by imparting energy to its planetary electrons, or decreased when one of its planetary electrons gives off energy, but that the energy imparted or released must always be an "allowed" amount. A wide variety of energy changes is possible, since the energy before and after the change can correspond with any two integral values of n in Equation (28), but the changes are nevertheless restricted to those corresponding to *integral* values of n.

The Relation of Atomic Structure to the Weight and Chemistry of the Various Kinds of Atoms.—There exist in nature more than 90 different kinds of atoms, each of which has its own set of

chemical properties. The 90-odd varieties are commonly listed in order of their atomic weight, beginning with hydrogen, the simplest and lightest, and ending with uranium, the heaviest and most complex. Each element in this listing is given a number; the atomic number of hydrogen is 1, that of uranium 92.

To explain these differences in weight and chemical properties in terms of the nucleus and the planetary electrons, it is assumed that the heavier atoms have large heavy nuclei and a large number of revolving electrons. Moseley showed that the

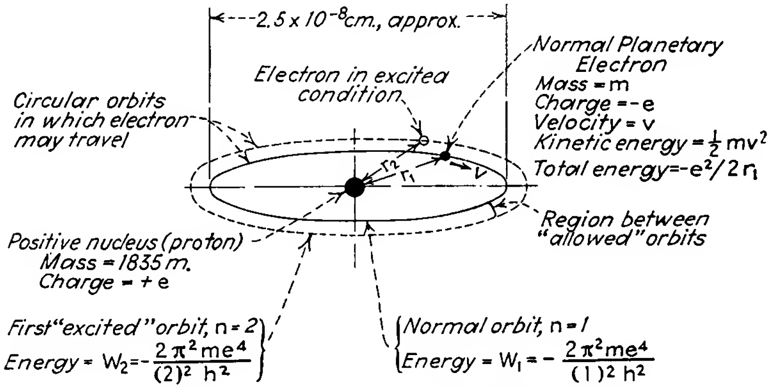


FIG. 41.—Energy relations in the hydrogen atom.

number of external revolving electrons in each neutral atom corresponds with the atomic number of that atom, and that the positive charge residing in the nucleus is equal to the atomic number times the unit positive charge.

In the simple lightweight atoms, such as hydrogen and helium, the combination of electrons and their arrangement are not complicated. In hydrogen, for example, the atom consists of a single positive charge, or proton, with one electron revolving about it. The energy relationships in this case are shown in Fig. 41. Helium has a nucleus consisting of two protons and two neutrons (a neutron is a chargeless body having the same weight as a proton) so that its weight is four times that of the hydrogen nucleus while its charge is only twice as great. The helium atom has two electrons revolving about the nucleus. Each of these electrons may have any allowed energy, according to Equation (28), and any corresponding allowed orbit. In the

normal condition of the atom, however, they inhabit the orbit corresponding to the value of $n = 1$. When energy is added to the helium atom, one or both electrons may jump to a new orbit having a larger radius, but the new orbit will always be one of the allowed orbits.

If we consider a heavy atom, such as mercury, we find a much more complicated state of affairs. Mercury has an atomic number of 80; consequently its nucleus contains 80 electron units of positive charge, which is balanced by the combined negative charge of 80 electrons revolving about it. These electrons are arranged in groups or "shells."[1] The outer shell, being the most exposed, is most susceptible to external influence. Therefore the electrons in the outer shell are those to be considered in explaining the action of a mercury atom in relation to external electrons which may collide with it or pass near by. The mercury atom is particularly important from a practical point of view, since mercury vapor is very commonly used as the conducting medium in gas-filled tubes.

Of equal importance with its electrical arrangement is the over-all mass of an atom, since this mass determines the rate at which a charged atom will accelerate in an electric field. The mass of the hydrogen atom may be determined by noting the deflection of a beam of charged hydrogen atoms (hydrogen nuclei, *i.e.*, protons) under a given electric field. By this means it was found that the mass of the hydrogen nucleus is approximately 1835 times as heavy as the electron. The atomic weight of mercury, determined by chemical means, is found to be 200 times that of hydrogen, so it weighs about 370,000 times as much as an electron. This fact is of importance in determining the current which a charged mercury atom (mercury ion) can carry in flowing between two electrodes in a tube.

Summary.—The neutral molecule of which gases and vapors are composed is made up of one or more neutral atoms. These atoms consist of a heavy positively charged nucleus, surrounded by a number of electrons equal to the atomic number of the element in question. The energy which these electrons may possess, and the radii of the corresponding orbits, are restricted to certain "allowed" values. Each electron may accept energy

[1] For a concise account of atomic chemistry see Structure of Atoms and Molecules, M. L. Huggins, *Elec. Eng.*, June, 1934.

from an external source, provided that the amount accepted is sufficient to raise the electron energy to another of the "allowed" values. If the energy is not sufficient to do this, then the energy is not accepted. Likewise an electron with more than its normal amount of energy may release energy, but only in such an amount that the left-over energy will be equal to an allowed energy value.

21. Electron-molecule Encounters. The Transfer of Energy from an External Electron to a Molecule.—With these facts in mind we may now consider the action of a free electron as it moves through a body of gas molecules. We consider only monatomic molecules, *i.e.*, molecules consisting of a single atom, since all the commonly used gases and vapors are of this type. Consider first a container filled with such a gas; at standard temperature (0°C.) and pressure (760 mm. Hg) there are 2.7×10^{19} molecules in each cubic centimeter. These molecules are moving, at random, in any possible direction and with any possible speed. This heterogeneous motion results in many collisions between molecules, but the collisions are not energetic enough to do any permanent damage to any molecule, so each molecule remains in its neutral uncharged state.

Now, if we place in this container two electrodes, one of which is an emitter of free electrons, and connect between the electrodes a battery with the negative terminal connected to the emitter, the free electrons will be attracted away from the emitter toward the near-by positive electrode. But the electrons cannot move far before they collide with one of the gas molecules which fill the space between the electrodes. Since there are 2.7×10^{19} molecules in a cubic centimeter, the average distance between the molecules is very small. There are $\sqrt[3]{2.7 \times 10^{19}}$ molecules along each edge of the cubic centimeter. The average distance between their centers then must be $\frac{1}{3} \times 10^{-6}$ cm. Since the molecules are small (about 10^{-8} cm. in diameter), and since the electron is smaller still (about 10^{-13} cm. in diameter), the electron may travel past many molecules before it collides with one. But at best it can travel only a very small fraction of a centimeter (about 10^{-5} cm. on the average) before it collides with a neutral molecule. The distance it moves is so small that it cannot acquire a high speed before the collision, hence its energy at the time of collision is small. Only if a very high voltage is applied

between the plates can the electron energy acquire high enough energy to disrupt the molecule.

But if the pressure inside the container is reduced, a very different state of affairs exists. Suppose the pressure is reduced to $\frac{1}{1000}$ of atmospheric pressure ($=0.76$ mm. Hg). Then the number of molecules per cubic centimeter is

$$\frac{27}{1000} \times 10^{18} = 27 \times 10^{15}.$$

There are therefore 3×10^5 molecules on each edge of the cubic centimeter, and their average separation is $\frac{1}{3} \times 10^{-5}$ cm. This separation is ten times as great as at atmospheric pressure, and the free path of the electron before collision is thereby greatly increased. Since the electron can move farther, the energy it acquires under the force of a given electric field is correspondingly greater, and its energy at the collision thus increases rapidly as the pressure of the gas is reduced. With a given voltage applied between the electrodes in the gas, therefore, the lower the pressure the more disruptive the collisions between the free electrons and the molecules. Likewise, with a given pressure, the higher the applied voltage, the more energetic the collisions. With a given pressure and voltage, the more closely spaced are the electrodes, the greater the electric field and the harder the electron-molecule encounters.

The average unobstructed path L of the impacting electron (its *mean free path*) may be calculated from the following:

$$L = \frac{4}{\pi D^2 n} \tag{29}$$

where L is the free path in centimeters, D is the diameter of the molecule in centimeters, and n is the number of molecules per cubic centimeter. The relationships between molecule density, electron and molecule diameters, and the resulting mean free path, together with the energy acquired by the electron in its free flight, are given in graphical form in Fig. 42.

Electron-molecule Energy Transfer.—By varying the pressure, the separation between the electrodes, and/or the voltage applied to them, therefore, it is possible to endow the free electron with almost any desired degree of energy by the time it strikes one of the gas molecules. The energy of the impacting electron at the time of collision is expressed in electron-volts (1 electron-

volt, the energy acquired by an electron passing through a voltage difference of 1 volt, is 1.6×10^{-12} erg). Suppose by varying the voltage between the electrodes immersed in low-pressure gas that the energy of the impacting electron is gradually

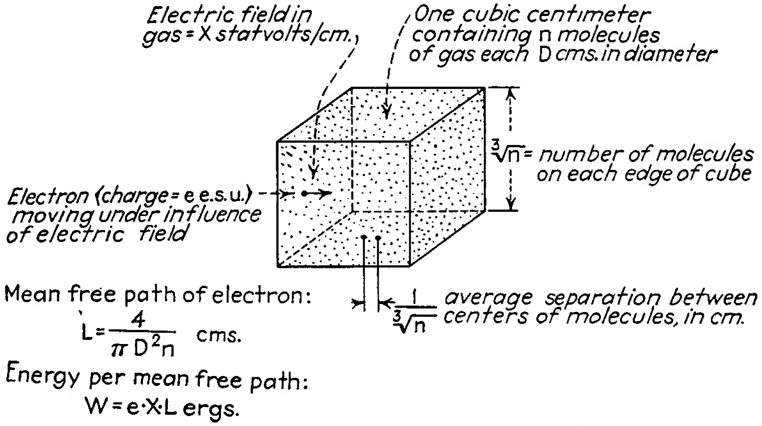


FIG. 42.—The energy attained by an electron moving in a gas depends on the mean free path between its collisions with the gas molecules.

increased. We now inquire how much of the energy is transferred from the impacting electron to the molecule (see Fig. 43).

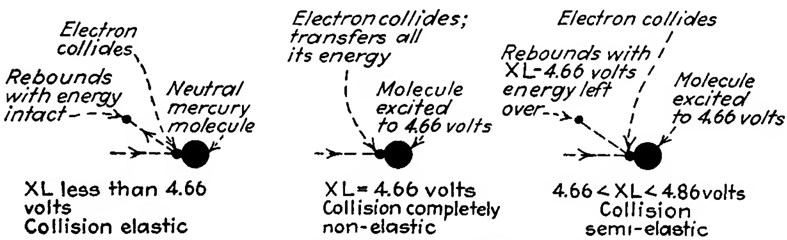


FIG. 43.—Three types of collision between an electron and a mercury atom, depending upon the energy of the electron.

If the electron energy is less than 2 volts, none of the energy is transferred. The collisions between electron and molecule are almost perfectly elastic. But as the energy of the electron is increased, a point is reached where the energy is transferred wholly from the electron to the molecule. At this point the impacting electron has just the right amount of energy to raise the energy level of one of the outer electrons in the gas molecule to the next allowed value. In sodium, this occurs when the

impacting electron energy is 2.10 volts; in mercury 4.66 volts; in helium it is 19.73 volts; argon, 11.57 volts; neon, 16.60 volts. These energy values are called the *first critical potentials* of the gases in question.

When a molecule has accepted the impacting electron's energy, it is said to be in an *excited condition.* But the molecule does not long remain in this condition. After about 10^{-8} sec., the extra energy it has absorbed is suddenly released, in the form of radiant light, the wave length of which depends on the amount of energy released. The molecule is thereby restored to its normal condition, and may repeat the performance if another electron

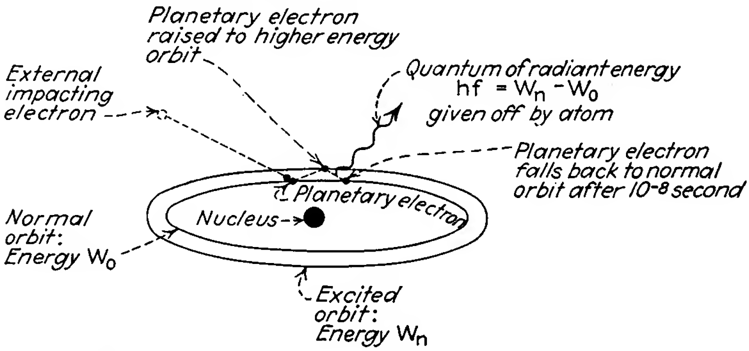


FIG. 44.—When an excited planetary electron returns to its normal orbit, a quantum of radiant energy is emitted by the atom.

bombards it with sufficient energy. This interesting behavior (see Fig. 44) is the explanation of the fact that the gas gives off light when the applied electric field is high enough and when the pressure is low enough.

If the impacting electron has a higher energy than the first critical potential, it may raise the molecule to a still higher allowed energy state. The molecule may then return to normal in one or more steps, jumping between successive allowed energy levels. The light output of an excited molecule is therefore usually very complex; only a small part of it is in the visible spectrum. The details of this effect are discussed at greater length in Chap. IX.

22. Ionization, the Removal of an Electron from Its Parent Atom.—The collisions thus far considered leave the molecule in an electrically neutral state. If the energy supplied by the impacting electron to the molecule is great enough, an outer

electron of the molecule can be made to leave the molecule altogether. This removal of the negative electron (see Fig. 45) leaves the molecule in a positively charged state. The molecule

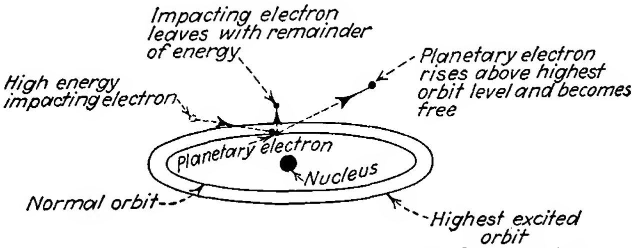


FIG. 45.—Ionization. When the impacting electron has sufficient energy, it releases a planetary electron.

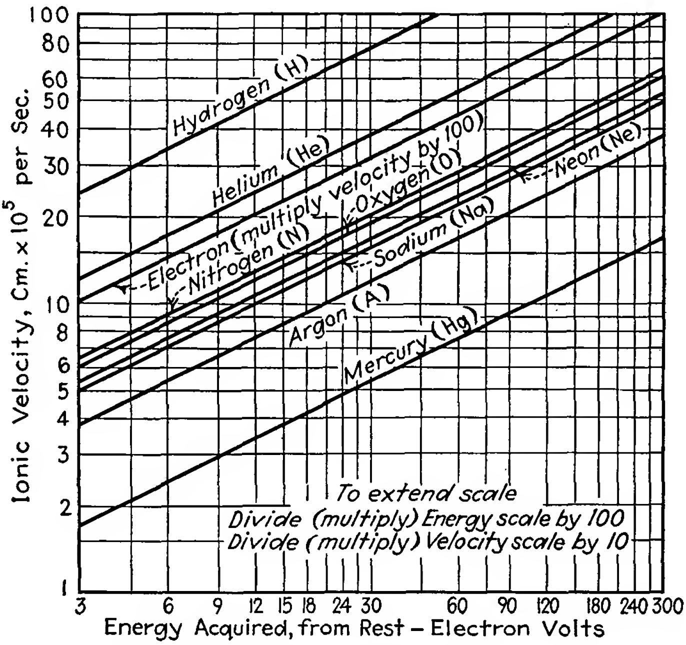


FIG. 46.—Velocities achieved by positive ions in terms of the voltage difference through which they pass.

is then called a positive ion, and the process by which it loses an electron is called ionization.

The energy level required for the removal of the electron has been measured for many gases and vapors. In mercury vapor, for example, it is found that if the impacting electron has an

energy of 10.38 volts the molecule with which it collides will become ionized. In argon the required energy is 15.69 volts; in neon, 21.47 volts; in helium, 24.48 volts.

The ionization of a gas molecule, since it is the production of a positively charged particle, is of the greatest significance in the electrical action of the gas. Since the ionized molecule is charged, it will move under the influence of an electric or magnetic field and will consequently carry a current. But the molecule is so heavy compared with the electron that its motion is much more sluggish (see Fig. 46), and the amount of current it can carry under a given field is comparatively small. To illustrate this fact, we may compute the current carried by a single singly-charged mercury ion (mass = 370,000 × electron mass) flowing between parallel plates separated by 0.3 cm. and connected to a 200-volt battery (compare with Prob. 5 at the end of Chap. II, page 32).

Problem 1. Given:

$e = 4.80 \times 10^{-10}$ e.s.u.

$m = 370{,}000 \times 0.91 \times 10^{-27} = 3.37 \times 10^{-22}$ g.

$s = 0.3$ cm.

$E = 200$ volts $= \frac{2}{3}$ statvolt.

To find: The force f on the mercury ion:

$$f = \frac{eE}{s} = \frac{4.80 \times 10^{-10} \times \frac{2}{3}}{0.3} = 10.6 \times 10^{-10} \text{ dynes.}$$

The acceleration a undergone by the ion:

$$a = \frac{f}{m} = \frac{10.6 \times 10^{-10}}{3.37 \times 10^{-22}} \text{ cm./sec./sec.}$$
$$= 3.16 \times 10^{12} \text{ cm./sec./sec.}$$

The time of flight of the ion:

$$t = \sqrt{\frac{2s}{a}} = \sqrt{\frac{0.6}{3.16 \times 10^{12}}} = 4.4 \times 10^{-7} \text{ sec.}$$

The average velocity (one-half final velocity)

$$v_{av} = \frac{v}{2} = \frac{at}{2} = \frac{3.16 \times 10^{12} \times 4.4 \times 10^{-7}}{2}$$
$$= 6.95 \times 10^{5} \text{ cm./sec.}$$

A current of 1 amp. is equal to 0.624×10^{19} electrons (or ions of equal charge) passing a given point in a second, or a single electron moving through 1 cm. at an average speed of 0.624×10^{19} cm./sec. The ion moving through 0.3 cm. at 6.95×10^{5} cm./sec. thus con-

stitutes a current flow of

$$i = \frac{(6.95/0.3) \times 10^5}{0.624 \times 10^{19}} = 3.7 \times 10^{-13} \text{ amp.}$$

The current carried by a single electron (mass 1/370,000 that of the mercury ion), under similar circumstances, is about 600 times as great, *i.e.*, 2.2×10^{-10} amp.

TABLE IV.—ATOMIC PROPERTIES OF GASES AND VAPORS PRESENT IN ELECTRON TUBES

Substance	Atomic number	Nuclear charge	Electrons per shell	Atomic weight	Nuclear mass	Nuclear charge-to-mass ratio	First resonance potential	Ionization potential
Hydrogen (H)...	1	$+e$	1	1.008	$1{,}830m$	$\frac{1}{1830} \times \frac{e}{m}$		13.53
Helium (He)....	2	$+2e$	2	4.002	$7{,}320m$	$\frac{1}{3660} \times \frac{e}{m}$	19.77*	24.48
Nitrogen (N)....	7	$+7e$	2,5	14.008	$25{,}620m$	$\frac{1}{3660} \times \frac{e}{m}$		14.48
Oxygen (O).....	8	$+8e$	2,6	16.000	$29{,}280m$	$\frac{1}{3660} \times \frac{e}{m}$		13.55
Neon (Ne)......	10	$+10e$	2,8	20.18	$36{,}600m$	$\frac{1}{3660} \times \frac{e}{m}$	16.4* 16.7	21.47
Sodium (Na)....	11	$+11e$	2,8,1	23.00	$42{,}090m$	$\frac{1}{3830} \times \frac{e}{m}$	2.10	5.12
Argon (A).......	18	$+18e$	2,8,8	39.94	$73{,}200m$	$\frac{1}{4070} \times \frac{e}{m}$	11.57	15.69
Mercury (Hg)...	80	$+80e$	2,8,18, 32,18,2	200.61	$366{,}000m$	$\frac{1}{4580} \times \frac{e}{m}$	4.66* 4.86	10.38

e = electron charge = 4.80×10^{-10} e.s.u. $\frac{e}{m}$ = electron charge-to-mass = 0.5×10^{18} e.s.u./g.
m = electron mass = 0.91×10^{-27} g.

* Metastable: no radiation results until subsequent transition to other than normal state.

Since the current carried by the ions is small, it is usually of no consequence in practical tubes. However, two other effects of the positive charge are of the greatest importance. The first is the ability of the positive ion to recombine with any near-by electron. A concentration of positive ions in any part of the tube can thereby remove from circulation an equal number of current-carrying electrons, if given the chance. The second effect is the ability of the positive field which surrounds each positive ion to nullify in part the negative field surrounding each near-by electron, and thereby greatly reduce the repelling forces between the electrons themselves. This modifies the net force acting on the electrons, to a degree, in fact, which may completely remove the interelectronic forces. The 3/2-power

law therefore does not apply since this law is based in part on the presence of repelling forces between the electrons. The current flow through a gas-filled space is thus not calculable in the same way as that through a vacuum. In fact, the current flow through a gas-filled space is almost impossible to calculate because of the large number of controlling factors present.

23. The Effects of Ionization in a Gas. The Self-maintained Gas Discharge.—When the energy of the free electrons in a gas is great enough to cause appreciable ionization of the gas molecules, *i.e.*, when the applied electric field is high enough in relation to the pressure, the gas *breaks down* and its resistance to electric-current flow becomes much smaller than it is in the normal un-ionized state. The increased current which flows is called a *gas discharge.* It is usually accompanied by a visible glow or arc. The gas discharge is a highly complicated phenomenon, and its importance in practical electronic tubes is very great. It exists in several forms, the Townsend discharge, the glow discharge, and the arc discharge, each of which has distinguishing features.

The Townsend discharge occurs when the degree of ionization is very small. Any body of neutral gas is partially ionized by the effect of cosmic rays and radioactive radiations which are present everywhere. The molecules ionized by these agents give rise to free electrons. These free electric charges are capable of carrying a small current, usually not more than a fraction of a microampere, between two electrodes immersed in the gas. If the container is covered with a thick sheath of lead, the cosmic and radioactive radiations are excluded, the ionization ceases, and the gas becomes an almost perfect insulator. The Townsend discharge is not self-maintaining and is of no practical importance.

If a high voltage difference is applied between the electrodes in the gas, however, the small number of free electrons originally present are given a high energy and produce new ions by colliding with neutral molecules. Each ionizing collision gives rise to a new free electron, which in turn gains energy from the electric field and collides with other neutral molecules, producing new ions and new free electrons. The process, being cumulative, soon produces a high concentration of positive ions and free electrons, both of which are available for carrying current.

There is, of course, always present the tendency of the ions and electrons to recombine into neutral molecules, but if the rate of ionization exceeds the rate of recombination, the number of ions and free electrons steadily increases.

As this building-up process continues, a point is reached where each free electron *regenerates* itself, and the discharge then becomes self-maintaining, *i.e.*, it maintains itself independent of all external influences except the applied electric field. Since this type of discharge depends on maintaining a supply of electrons in excess of those recombining with ions, it is necessary that the extra supply of electrons come not from neutral gas molecules but from some other part of the tube, usually from the electrodes.

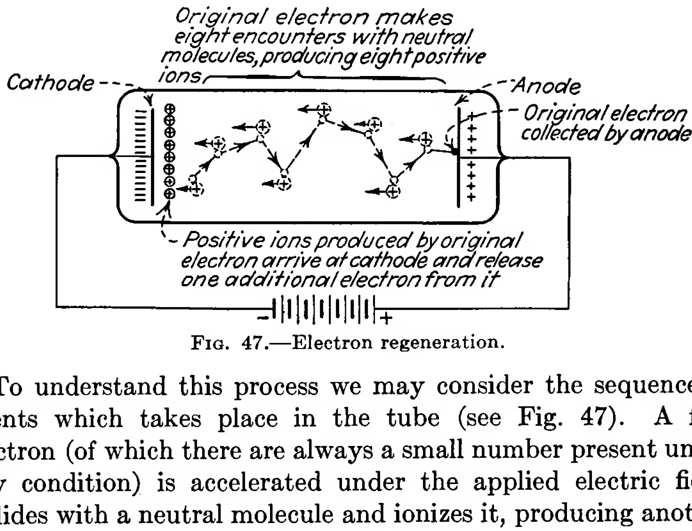


FIG. 47.—Electron regeneration.

To understand this process we may consider the sequence of events which takes place in the tube (see Fig. 47). A free electron (of which there are always a small number present under any condition) is accelerated under the applied electric field, collides with a neutral molecule and ionizes it, producing another free electron and an ion. The free electron accelerates in the field and produces still another free electron and a new ion, and so on. The ions produced in this process are accelerated, but at a slower rate and in the opposite direction from the electrons. Eventually these ions are attracted to the near vicinity of the negative electrode. Then, if the total number of ions produced by the original free electron can by some means liberate another free electron from the negative electrode, the original electron regenerates itself, and the supply of free electrons is maintained regardless of the tendency to recombine with ions.

The concentration of positive ions near the negative electrode can free an electron by any one of several methods. The ions may bombard the electrode and free electrons by secondary emission. The bombardment may heat the electrode sufficiently to induce thermionic emission. Or the light present, particularly the ultraviolet light, may induce photoelectric emission. The process is inefficient, and it is necessary that the original electron have a large ionic progeny so that all the ions attributable to it will produce another electron from the electrode. But when this happens, the gas breaks down and the discharge thereafter maintains itself.

If the current flow through the gas is maintained at moderate levels, of the order of milliamperes (*i.e.*, if the resistance of the external circuit prohibits a larger current flow than this), then the discharge takes the form of a glow discharge. This type of discharge is characterized by a fairly high internal resistance in the gas, and the voltage drop across the electrodes is high, of the order of several hundred volts. But if a higher current flow is allowed to pass, the ionization becomes correspondingly more vigorous, and the induced emission of electrons at the negative electrode is heavy. The discharge then becomes an arc discharge. The temperature inside the gas increases greatly, and at the same time the voltage drop across the electrodes decreases markedly, to about 10 to 25 volts. The current in the arc discharge is limited only by the resistance of the external circuit; currents of many thousands of amperes may be carried, representing the extraction of enormous numbers of free electrons from the negative electrode.

In many gas-filled electron tubes the production of electrons at the negative electrode is aided by making the electrode a thermionic emitter, and heating it from an external power source. The arc so produced is sometimes called an artificial arc, but its characteristics are exactly the same as the natural arc in which an untreated metallic electrode is used.

The voltage at which the breakdown begins depends on the pressure of the gas. As the pressure is reduced, the free path of the electrons between collisions is increased, as already stated. The energy picked up by the electron before each collision occurs is thus greater and its ability to ionize the molecule is enhanced. Therefore, the lower the pressure, the lower

the applied voltage at which the breakdown occurs. At very low pressures, however, the number of available molecules is reduced faster than the energy of the electrons is increased, so the chance of the electron producing ions is reduced by sheer lack of numbers. Hence at very low pressures, the voltage required for breakdown increases as the pressure decreases. Midway between these two opposed tendencies is a *minimum breakdown voltage*. A typical breakdown curve illustrating this effect is shown in Fig. 48.

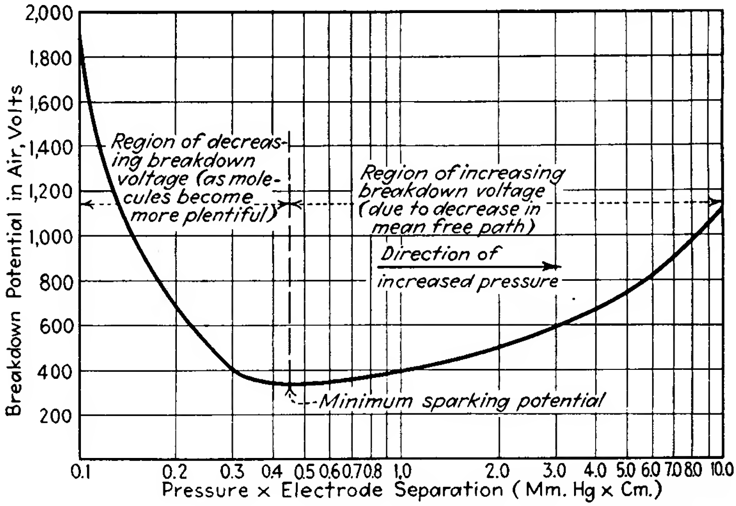


FIG. 48.—Breakdown voltage of air between plane electrodes. (*After Schumann.*)

At any one pressure, the tendency of electrons to ionize depends on the applied voltage divided by the separation between the electrodes. If the gas breaks down with one value of applied voltage with a given separation of electrodes, it will break down with the same applied voltage if the gas pressure is halved and the separation between electrodes is doubled. This is Paschen's law, one of the few dependable quantitative relations available in the entire theory of the gas discharge. The actual value of the voltage required for breakdown depends on the separation of the electrodes, since this separation determines the amount of voltage drop available in the free path of each electron.

24. The Control of Electron Currents in Gases and Vapors.—The control of electron currents in a gas or vapor is very different

from that in a vacuum. The amount of current flow depends, in the first place, upon a great many factors: the type of gas, its pressure and chemical purity, the composition, shape, and surface treatment of the electrodes and their separation, the value of the applied voltage and the frequency with which it changes polarity, and the current-supplying ability of the power source. In the face of this array of independent variables, about all that can be said is that when sufficient voltage is applied between electrodes in a gas, a current will flow between them, and that when the voltage is removed, the current flow ceases. The amount of current flow can be predicted only approximately, even if all the controlling factors are specified. It is known

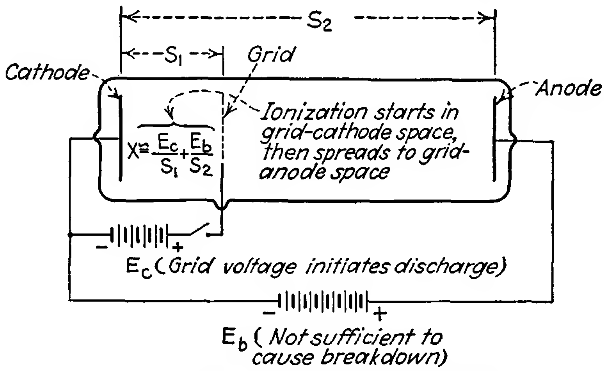


FIG. 49.—Grid control of the gas discharge.

also that the start of the current flow and its cessation are not instantaneous. The time required to initiate the breakdown is called the *ionization time,* and that to remove all traces of the ionization after the voltage is removed is called the *de-ionization time.* In practical cases these times vary from one to a thousand microseconds (millionths of a second). This delay is a severe limitation on the utility of gas-filled electron tubes to handle rapidly varying currents. It is treated in more detail in Chap. VII.

The amount of current flow in a given tube can be controlled by changing the gas pressure, an impractical method, or by changing the resistance of the external circuit. More practical is the use of grid control, *i.e.*, the insertion of a grid structure between the current-carrying electrodes (see Fig. 49). By applying a voltage difference between this grid and the negative

electrode, a limited degree of current control can be obtained. It is found that the breakdown of the gas can be initiated by making the grid sufficiently positive with respect to the negative electrode, since the field between them is thereby increased and the energy of the free electrons increased to the ionization value. But once the discharge has begun it is impossible to modify the current further by changing the voltage applied to the grid. This limited behavior, in direct contrast with the grid control of electron currents in a vacuum, is explained by the fact that the grid becomes coated with a sheath of positive ions, which are attracted to it. The thickness of this sheath increases and decreases as the grid voltage is changed, and thereby neutralizes the field which would otherwise surround the grid. The grid can thus be used only to start the discharge, but not to change the current flow in any way thereafter. This limitation has not prevented the application of grid control in a wide variety of useful gas-filled electron tubes. These tubes and their dependence on gas-discharge effects are discussed in Chap. VII.

Problems

1. Calculate the first five "allowed" energy levels of the single electron in a hydrogen atom. Express the results in ergs and in equivalent volts.

2. Calculate the radii of the electron orbits corresponding to the energy levels in Prob. 1.

3. Calculate the acceleration undergone by a hydrogen ion between parallel plates separated by 1.5 cm. and connected to a 300-volt battery. Compare this acceleration with that undergone by an electron and that by a mercury ion, in the same circumstances.

4. Calculate the average separation between the centers of the molecules in mercury vapor at a pressure of one millionth of an atmosphere. What pressure corresponds to a separation of 25 cm.?

5. Calculate the mean free path of an electron in mercury vapor at 0.1 mm. pressure. (Diameter of Hg ion 10^{-8} cm.)

6. Ionization in H_2 gas between parallel plates generates 10^{15} ions and 10^{15} electrons per cubic centimeter per second. If 50 per cent of the ions and electrons recombine and if the applied electric field removes the remaining ions and electrons as fast as they are generated, what is the total current carried between the plates? Plate area, 100 sq. cm. Plate separation 5 cm.

7. In a mercury-vapor tube each free electron liberates on the average 15 ions before it recombines. If 50 per cent of these ions recombine by the time they reach the cathode, and if the efficiency of electron production at the cathode is 5 ions per electron, will electron regeneration occur? What efficiency of ion production by the original electron is required to start regeneration, all other factors remaining the same?

Bibliography

HULL, G. F.: "An Elementary Survey of Modern Physics," Chaps. 1, 3, 4, 6, and 9, The Macmillan Company, New York, 1936.

DARROW, K. K.: "Electrical Phenomena in Gases," Williams and Wilkins Company, Baltimore, 1932.

THOMSON, SIR J. J., and G. P.: "Conduction of Electricity through Gases," University Press, Cambridge, 1928–1932.

LOEB, L. B.: "Kinetic Theory of Gases," McGraw-Hill Book Company, Inc., New York, 1927.

TOWNSEND, J. S.: "Electricity in Gases," Clarendon Press, Oxford, 1915; "Motions of Electrons in Gases," Clarendon Press, Oxford, 1925.

RUARK and UREY: "Atoms, Molecules, and Quanta," McGraw-Hill Book Company, Inc., New York, 1930.

Physics Department, University of Pittsburgh, "An Outline of Atomic Physics," 2d ed., John Wiley and Sons, Inc., New York, 1937.

A series of articles in *Electrical Engineering*, Electric Discharges in Gases: Part I, Lewi Tonks, February, 1934; Part II, K. K. Darrow, March, 1934; Part III, J. Slepian and R. C. Mason, April, 1934. Also A. W. Hull, Fundamental Properties of Mercury Vapor and Monoatomic Gases, November, 1934.

PART II

ELECTRON TUBES

Their Constructions, Principles of Operation, and Characteristics

CHAPTER VI

THERMIONIC VACUUM TUBES

Introduction.—The principles of physical electronics with which we have been concerned in the foregoing chapters are applied to the design and construction of all classes of electron tubes. In the present chapter, the *thermionic vacuum tubes* are considered. The name indicates the two distinguishing characteristics of this group: They obtain electron emission by thermionic means, and the emitted electrons flow through a vacuum, *i.e.*, the tubes are substantially gas-free.

Thermionic vacuum tubes are the oldest of the electron-tube family, and—perhaps because they are oldest—they have been applied to the widest variety of uses. They are completely indispensable to radio transmission and reception and to long-distance electrical communication in all its forms. Since they are the only true amplifying tubes,[1] they are at the root of nearly all scientific electronic applications, and many industrial uses as well. The thermionic tubes are important not only from this practical point of view but also from the standpoint of the development of electronic principles, since the constructions used in them are applied also in gas-filled tubes and to some extent in photosensitive tubes. Accordingly the study of the thermionic vacuum tube is the basis of all practical electronic engineering.

25. Thermionic-tube Classifications: Diodes, Triodes, Tetrodes, Pentodes, Multipurpose.—There are two elements common to all thermionic tubes, the electron emitter or *cathode*, and the electron collector or *anode*. If only these two elements are present, the tube is called a *diode* (from *di* = two and *ode*

[1] Recently a new form of amplifier, the electron multiplier tube (page 214), has appeared which does not depend on thermionic emission, but its application, thus far, is limited to current amplification of electron beams initiated by the action of light on photosensitive surfaces and to the production of self-sustained oscillations, neither of which is "amplification" in the sense here intended.

from electrode). In the diode, the electron current passes from the cathode emitter through the vacuum to the anode collector. Since the anode does not ordinarily emit electrons, no electron current can be carried in the reverse direction, from anode to cathode. The diode's ability to carry current in one direction fits it only for use as a rectifier of alternating currents. In one form or another this rectification action is the sole practical use of diode tubes.

If a grid structure is placed between the cathode and the anode, the number of electrodes present is three, and the tube is called a triode. The presence of the grid enormously increases the usefulness of the tube. By changing the voltage applied between the grid and the cathode, the current which flows between cathode and anode can be made to change practically instantaneously over a wide range. This control over an electric current by an entirely unrelated voltage is the basic action of the triode tube. It makes possible amplification, frequency conversion, modulation, demodulation, oscillation, and controlled rectification, which include every electronic function except photosensitivity. Because of its extraordinary versatility, the triode, and its cousins the tetrode and the pentode, outnumber all other electron tubes in service by at least 10 to 1.

Under certain circumstances the action of the triode is limited by the tendency of the anode-cathode circuit to react back on the grid-cathode circuit, so that the electron current is controlled not by an entirely unrelated voltage but by a voltage partly under the control of the current itself. To divorce this relationship, another grid (called a screen grid) is placed in the triode tube, bringing the number of electrodes to four, hence the tube is called a tetrode. The tetrode is extraordinarily sensitive, *i.e.*, changes in the voltage applied to its grid are much more effective in changing the electron current than similar changes in voltage applied to the anode. It is accordingly used as an amplifier of voltage changes. Except for the presence of the extra grid the tetrode is in no way different from the triode, so far as its essential action is concerned.

Still another grid is placed in the fourth class of tube, the pentode, for the purpose of suppressing secondary emission from the anode, which subtracts from the desired electron flow from cathode to anode. The pentode, like the tetrode, is

thus really a triode in disguise. It operates in the same manner as a triode, so far as the external circuit is concerned, but its two extra grids make its action far more effective for certain purposes. Nearly all radio tubes used in modern radio sets are pentodes, because of their ability to amplify both voltage and power to high levels within the confines of a single tube.

Many tubes in the thermionic family are combinations of diodes with triodes, or diodes with pentodes, etc., the two sets of elements being placed within a single tube purely for reasons of convenience and economy. The names of these multipurpose

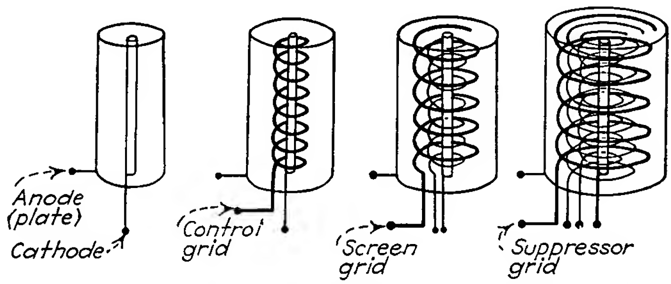


Fig. 50.—Elemental forms of thermionic vacuum tubes; left to right: diode, triode, tetrode, and pentode.

tubes are combinations of their component names. The *duo-diode-pentode* used in many modern superheterodyne receivers, for example, contains two diodes and one pentode, all operating to a common cathode, but otherwise separate. The combination of tubes within one glass or metal envelope is dictated by the fact that the included tubes are always used in combination with each other.[1]

A specialized type of multipurpose tube is the *pentagrid converter* used for frequency-conversion service in superheterodyne receivers. This tube, sometimes called a heptode, contains five grids in addition to the cathode and anode. The grids serve the purpose of mixing two alternating currents, one the incoming radio signal and the other from a local oscillator, and the tube thus serves the combined actions of oscillator, mixer, and amplifier.

[1] For a full description of these multipurpose tubes as well as the simpler types see "The Radiotron Manual" and "Tube Handbook," issued by the RCA Radiotron Division of the RCA Manufacturing Company.

Finally there are many specialized tubes which are properly called thermionic vacuum tubes, but which bear little resemblance to the other members. The cathode-ray tube, discussed in Chap. IV (page 65) is an example. These members of the family are treated in a separate chapter (Chap. X, page 200).

26. Thermionic-cathode Structures: Construction, Operation, and Ratings.—The negative electrode, or cathode, of thermionic tubes is designed on the basic theory described in Chap. III, Sec. 9, page 38. There are three types of material used: pure metals, of which tungsten is the outstanding example, contaminated metals (thoriated tungsten) and metallic oxides, of which barium-strontium oxide is the most widely used.[1]

Pure-metal Cathodes.—Pure tungsten (work function 4.52 volts, melting point 3660°K.) is used as a thermionic-cathode material because it can be operated at high temperature. Its work-function voltage is not particularly low, so it is not particularly well suited for the purpose, but at the high operating temperatures it can withstand (up to 2700°K.), it gives sufficient emission for use in many types of tubes. Tungsten is a rugged material from the thermionic standpoint, and it will emit steadily under a wide variety of conditions, including the presence of a high electric field strength at its surface. It is accordingly used in high-power transmitting and rectifying tubes, where the applied voltage may be as high as 20,000 or 30,000 volts, and in X-ray tubes where the applied voltage often may approach a million volts. For tubes operating at lower applied voltages tungsten has no advantage over the metallic-oxide emitters from the standpoint of operating ruggedness, and its considerably lower efficiency of emission puts it out of the running.

Tungsten cathodes used in practice usually take the form of a wire filament, either straight wire or wire coiled into a spiral shape. Attempts to use a flat ribbon (which has a large surface area and hence greater total emission for a given weight of wire) have not been successful because of the mechanical fragility of the metal. Tungsten is not normally ductile and must be specially treated before it can be drawn into wire or ribbon. When brought to operating temperature, the metal recrystallizes

[1] A more detailed description of thermionic-cathode structures is given in L. R. Koller, "Physics of Electron Tubes," 2d ed., McGraw-Hill Book Company, Inc., New York, 1937, Chaps. II and III.

and loses its ductility. Tubes employing tungsten cathodes must accordingly be handled carefully whenever moved.

Thoriated tungsten, an impure form of the metal which has a much lower work function and an operating temperature of about 1900°K., is used now only in medium and low-power transmitting tubes where oxide emitters are not suitable. In high-voltage tubes, thoriated tungsten is far less rugged than pure tungsten, and hence it is not used.

Other pure metals suitable for thermionic-cathode use are tantalum and molybdenum, both of which have high melting points and middle-range work-function values. Neither is used to any extent in electron tubes, except for experimental purposes, because they offer no great advantage over pure tungsten.

Metallic-oxide Emitters.—Wehnelt[1] in 1904 discovered that a platinum strip covered with oxides of the alkaline metals strontium, barium, and calcium would produce large quantities of free electrons at low temperatures. For many years thereafter the oxide emitter was neglected, but in 1920 it was resurrected by the commercial tube companies. Today it occupies the most important position of all the thermionic-emission materials, forming the cathode surface of all modern receiving tubes (of which many millions are produced annually) and of many larger tubes as well.

The work-function values of the oxides of barium and strontium are very low, but their exact values are not definite. Usually the emitting material is a mixture of barium and strontium oxides, 50 per cent of each by volume. The measured work functions for such a combination vary from about 0.5 to more than 1.5 volts. These low values permit operation at a dull red heat, from 900 to 1200°K., with emission as high as 500 ma./sq. cm.

The commonly accepted explanation[2] of the action of the barium-strontium oxide is as follows: During the manufacture of the tube, a voltage is applied between the emitter cathode and the anode. This voltage gives rise to an electric field within the oxide, which is thereby electrolyzed. The metallic (positive) ions of barium and strontium migrate under the

[1] WEHNELT, *Ann. Physik*, **14**, 425 (1904); *Phil. Mag.*, **10**, 80 (1905).

[2] CHAFFEE, E. L., "Theory of Thermionic Vacuum Tubes," McGraw-Hill Book Company, Inc., New York, 1933, pp. 102–114.

influence of the field to the negative metal base of the emitter, while the oxygen travels to the outer surface. Thereafter the metallic barium and strontium migrate, by diffusion action, to the outer layer and there form a very thin layer of metallic barium and strontium. The electron emission comes from this layer, primarily from the barium, which as a metal has a low work function, and which has an even lower value in thin layers.

There are three types of construction employed in the barium-strontium oxide cathodes, as shown in Fig. 51: the coated filament, the coated ribbon, and the indirectly heated structure. The barium and strontium, in the form of carbonates mixed

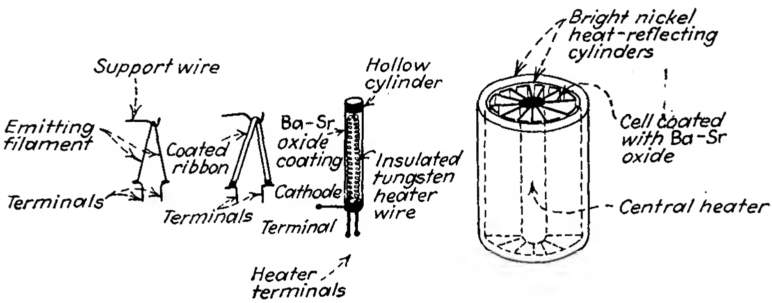


FIG. 51.—Thermionic-cathode structures. Left to right: wire filament, coated ribbon, indirectly heated cylinder, indirectly heated cellular structure.

with water or some organic binder, are applied as a liquid paste to a metal base, by either spraying, dipping, or painting. The composition of the metal base has an important bearing on the emitting ability of the surface; nickel and konel metal, an alloy of nickel, cobalt, iron, and titanium, are widely used.

The simplest construction is a wire of circular cross section coated with a thin layer of the barium-strontium carbonates. When a large amount of emission is required, a larger wire is necessary to increase the surface area. But the larger wire is costly, both in material and in the heating power required to bring it to operating temperature. To increase the surface area with a given weight of material, the coated ribbon filament is used. But the area cannot be increased very greatly by the use of thin ribbons before the ribbon becomes so thin that it loses its mechanical strength.

Further increases in emission with a given amount of base metal and heating power are obtained by the use of the *indirect-*

heater construction. In this type of emitter, the cathode coating is applied to a hollow structure, inside which is an insulated wire of tungsten, which heats the outer shell by radiation and conduction. The shell, which is usually cylindrical in shape, may have a very large area and may be very thin so that its thermal inertia is small. The tungsten wire has a high specific resistance and hence attains high heat with comparatively small current flow. Furthermore the size of tungsten wire may be properly proportioned to its length to give the maximum heat output in relation to the current flow. This construction is very flexible and is used in a wide variety of tubes, both small and large. It has a very important electrical advantage in the fact that the cathode emitting surface is isolated electrically from the heating circuit. The tungsten wire may accordingly be supplied with alternating current without the variations in the resulting alternating-current field having any effect on the emitted current.

The preparation of the oxide surfaces is accomplished by a combination of thermal and electrical treatment. The cathode, in filament, ribbon, or indirectly heated form, with the carbonate coating, is placed in the tube, after drying in carbon dioxide gas. The tube is then placed on the exhaust pump and the pressure in the tube reduced while current is applied to heat the cathode well above the operating temperature. This heat drives off the carbon and oxygen present in the carbonate coating, leaving barium and strontium oxides. The carbon dioxide gas formed in this process is removed by the pumps. A moderately high voltage (about 200 volts) is then applied between the anode of the tube and the cathode. This causes the electrolysis of the coating during which the layer of metallic barium and strontium gradually forms on the outer surface of the coating. During this time the current between cathode and anode gradually increases to the desired operating value. Gas given off by the coating, including oxygen from the electrolysis, is removed by the pumps. After this activation process the coating is ready for use.

Emission Efficiency of Thermionic Cathodes.—The engineering figure-of-merit, by which the thermionic emission properties of various substances and constructions may be compared is their *emission efficiency.* Emission efficiency expresses the total

(saturation) emission in milliamperes developed by a cathode, per watt of heating power supplied to it. The efficiency depends, therefore, not only on the work function and operating temperature of the surface but also on the thermal efficiency of the heating system. Some constructions conserve the heat supply much better than do others, and these constructions therefore have higher efficiency, other factors being equal.

Pure tungsten, in the form of wire, has an operating efficiency at rated temperature of about 5 ma./watt. Large heavy filaments may exceed this value somewhat, but in general the values are the lowest of all the practical emitters. Thoriated tungsten, having a lower work function, has an efficiency of about 25 ma./watt. Oxide-coated cathodes have a wide range of efficiencies, since the work function of the surface depends so much upon the treatment it receives during its preparation and activation and since a wide variety of heat-conserving constructions can be used. The simple coated wire has an efficiency of about 50 ma./watt. The cylindrical indirectly heated construction, commonly used in receiving tubes, supplies about 100 ma./watt, while the involved honeycombed constructions used in gas-filled rectifiers (unsuited to vacuum tubes because of space-charge limitations) achieve an emission of 1000 ma./watt.

These values are, of course, approximate averages only. It is usually necessary to take precautions in comparing different constructions to be sure that they are compared on an even footing. The input power, in particular, should be reduced to a unity ratio for the two samples, since the efficiency changes with the input power. Typical power-emission characteristics of tungsten, thoriated tungsten, and barium-strontium oxide are given in Fig. 52.[1]

The Saturation Emission of Cathodes.—The maximum current which can be carried through any thermionic tube depends, of course, on the ability of the cathode to supply the electrons which constitute the current flow. This ability is expressed by the Richardson emission equation [Equation (11), page 39]. The emission current given by this equation is, therefore, the maximum current obtainable from the cathode. In pure-metal emitters this limitation applies directly, but in the oxide-coated cathodes a rather peculiar behavior is exhibited. In the first

[1] DUSHMAN, Electron Emission, *Elec. Eng.* (July, 1934).

place, it is hard to predict the emission of oxide coatings in terms of the Richardson equation, since the work-function values and the values of the constant A are so dependent on the preparation and life history of the cathode. In the second place, it is found that as higher and higher voltages are applied between cathode and anode, the current flow increases almost without limit, far

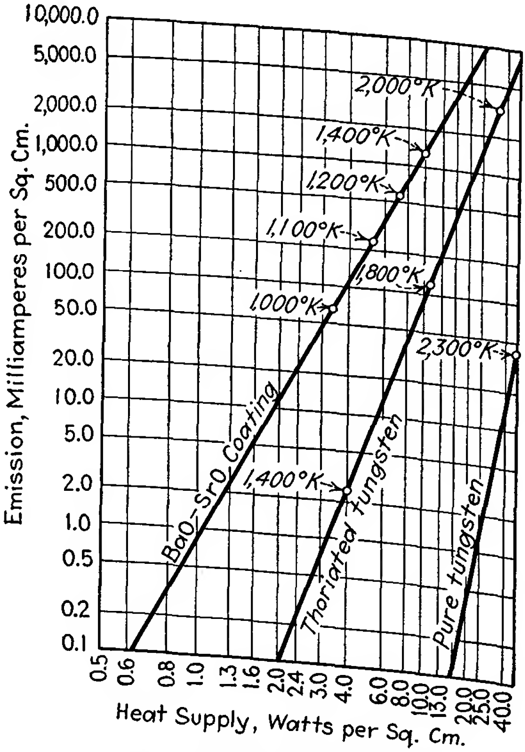


FIG. 52.—Emission *vs.* heat supply.

exceeding the saturation limit indicated by the equation. The explanation is that the increasing electric field at the surface gives rise to a further electrolytic action within the coating, supplying new metallic barium from which more emission is available. Consequently the maximum electron current obtainable from an oxide emitter is usually very large, far above the safe operating value. The value indicated by the Richardson equation is therefore not the maximum current but may be taken

as the safe operating limit which will give a satisfactory length of operating life.

The saturation limit of any cathode is also increased by the Schottky[1] effect, due to the lowering of the work function of the surface because of the presence of the attracting electric field at the surface of the emitter. This effect is usually small; it can be calculated from the Schottky equation:

$$I = I_s\epsilon^{\frac{4.39X^{1/2}}{T}} \tag{30}$$

where I is the emitted current, I_s the current predicted by the Richardson equation, X the value of the electric field in e.s.u. at the surface and T the temperature in degrees Kelvin.

27. The Diode and Its Characteristics.—In the diode tube, the cathode is surrounded by the electron-collecting anode, or *plate*. When a battery is connected between the cathode and plate, an electron current will flow between them, provided that the plate is connected to the *positive* terminal of the battery. The amount of current that flows depends, for any given tube, on the voltage difference between the cathode and plate, and also upon the temperature of the cathode. When the cathode temperature is fixed at its rated value, the current depends on the voltage applied. Consequently it is possible to express the relationship between current and voltage by a simple plot of electron current against applied voltage. Such a plot is called the *plate-voltage plate-current characteristic* or E_b-I_b *curve*.

Consider the circuit shown in Fig. 53, in which a diode tube is connected in series with a variable battery. A milliammeter measures the current flow issuing from the plate terminal, while a voltmeter measures the voltage applied between cathode and plate. If the applied voltage is increased step by step, and the current corresponding to each step recorded, the current will be found to increase in proportion to the $\frac{3}{2}$ power of the applied voltage. This follows from the $\frac{3}{2}$-power law [Equations (22), (23), and (24), page 58] which gives the current flow between electrodes when the negative electrode is emitting electrons.

As the applied voltage is further increased, a point is reached where the current flow no longer increases, or at best increases very slowly. At this point the saturation current is drawn from

[1] SCHOTTKY, *Physik. Z.*, **15**, 872, (1914).

the cathode, *i.e.*, all electrons liberated at the cathode are drawn to the plate. In tungsten-cathode tubes this saturation effect is quite pronounced, but in metallic-oxide cathodes less so, for the reasons given at the end of Sec. 26. The voltage at which this saturation sets in, depends, of course, on the emission of the cathode, which in turn depends upon the temperature of the

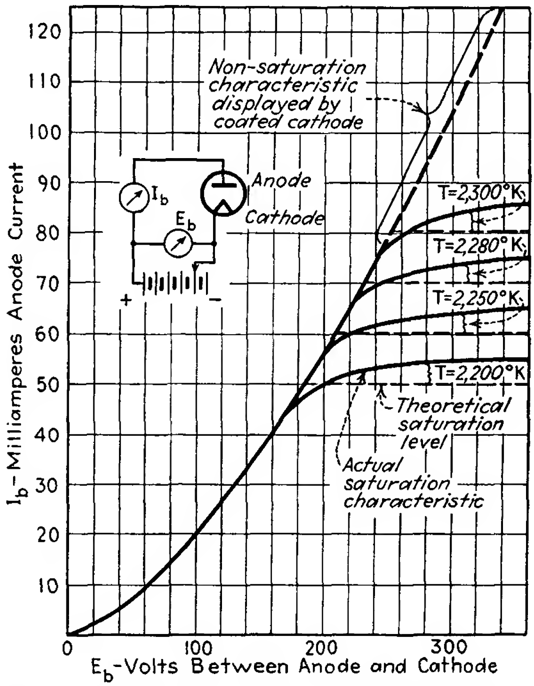


FIG. 53.—I_b-E_b curve family of a diode. The circuit shows the manner in which the data in the curves are measured.

cathode. Therefore by increasing the temperature of the cathode (by increasing the heater or filament current) the saturation level may be raised. Consequently the current-voltage characteristic of a diode consists of a family of curves (Fig. 53), each member of which corresponds to a definite filament temperature. The degree of saturation exhibited by the diode depends upon the type of emitter employed; in normal operation the saturation value is a higher value of current than can be provided by the cathode without loss of operating life.

The design of a diode tube begins with the cathode structure which must meet given emission-efficiency requirements and which must provide a saturation current considerably larger than the desired operating current. When a suitable emitter has been provided, the plate structure must then be designed to collect the electron current. This design involves the area and thickness of the plate material (since it must stand the kinetic energy of the bombarding electrons), and it also involves the shape of the plate and its separation from the cathode.

To illustrate the design of a diode tube, consider the following problem (Fig. 54):

Problem 1. Design a diode tube fulfilling the following requirements:

Cathode:

Emission efficiency, 100 ma./watt.
Saturation current, 200 ma.
Operating temperature, normal for type of cathode.

Plate:

Maximum voltage 300.
Maximum power dissipation at plate, 50 watts.
50 ma. emission at 150 volts.

An indirectly heated cathode of the oxide-coated type is called for by the emission-efficiency requirement. The normal operating temperature of this type is 1100°K. Using the Richardson equation [Equation (11)] and substituting the values $A = 10^{-2}$, $\phi_0 = 1.04$ volts, and $T = 1100$, the saturation emission is found to be 224 ma./sq. cm. Therefore a total cathode area of $^{200}\!/_{224}$ sq. cm. = 0.9 sq. cm. is required to fulfill the saturation requirement. This can be met by a cylindrical cathode 3 cm. long and 0.095 cm. in diameter.

To find the diameter of the cylindrical plate, use Equation (24),

$$I = \frac{14.68 \times 10^{-6} E^{3/2}}{r\beta^2} \text{ amp./cm. length.} \tag{24}$$

The current requirement is 50 ma., total (16.6 ma./cm. length) at 150 volts. Assuming a value of $\beta^2 = 1$ as a first approximation, substitute these values and solve for r:

$$r = \frac{14.68 \times 10^{-6}(150)^{3/2}}{(16.6 \times 10^{-3}) \times 1} = \frac{14.68 \times 1.84}{16.6} = 1.63 \text{ cm.}$$

Since this radius is 34 times that of the cathode, the proper value of β^2 is 1.09 (from Fig. 30), so that the assumed value of 1 introduces an error of 9 per cent. The proper radius of the plate is then 1.5 cm.

The current-voltage curve is then plotted for voltage values of 50, 100, 150, 200, 250, and 300, either by using Equation (24) for

each computation or by remembering that the current increases and decreases from 50 ma. as the $\frac{3}{2}$ power of the voltage. The current at 300 volts is $(300/150)^{3/2} = 2.83$ times that at 150 volts, *i.e.*,

$$2.83 \times 50 = 142 \text{ ma.},$$

while at 75 volts it is $50/2.83 = 17.7$ ma. The plot of these and the intermediate values is shown in Fig. 54. At the maximum

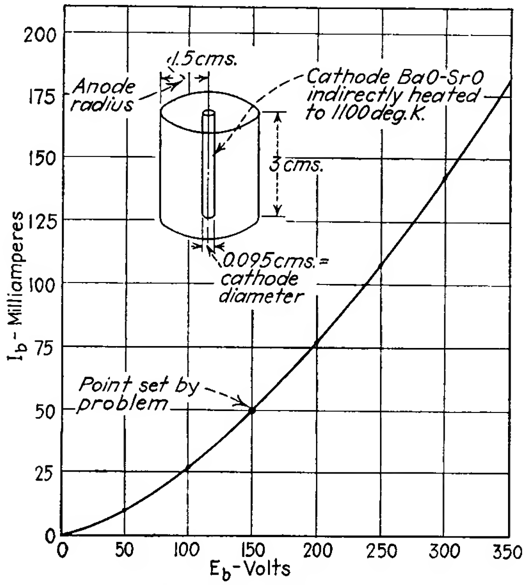


FIG. 54.—The design of a diode tube (Prob. 1).

voltage the product of the current flow through the tube times the applied voltage is $300 \times 0.140 = 42$ watts. Since this power is dissipated in heating the plate, the plate must be rugged enough to stand this heat dissipation. Exact calculations of thermal behavior are practically impossible, so this latter requirement is satisfied by reference to experience.

The Typical Diode-connection Diagram.—The diode is used almost exclusively in connection with an alternating voltage supply. This connection is shown in Fig. 55. The alternating voltage applied causes the plate to assume alternately a positive and negative sign with respect to the cathode; the current flows only during the positive periods. This is shown graphically in Fig. 55. The current flows through the tube only in one

direction, and only during the half cycles of the applied alternating voltage which make the plate positive. The tube therefore carries a unidirectional pulsating current or *rectified current* which has a host of special uses, discussed at length in Chap. XII. The form of each pulse of current, it will be noted, does not have the same *shape* as the half-cycle pulse of positive voltage which produces it, since the current-voltage curve is not a straight line. In detection action (see page 290), the nonlinearity of the rectification characteristic is of considerable importance.

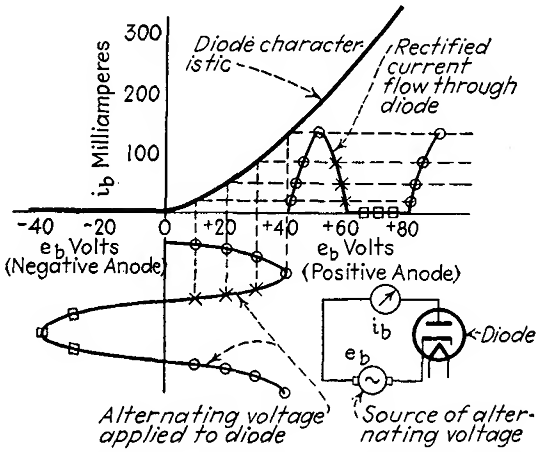


FIG. 55.—Use of the diode characteristic in determining current flow when alternating voltage is applied. The circles, squares, and crosses mark corresponding points on the voltage and current curves.

Before leaving the diode, it is necessary to consider what happens to the emitted electrons at the cathode when they are not wholly collected at the anode, *i.e.*, when the anode voltage is less than the value necessary for saturation. In this case the full quota of electrons predicted by the Richardson equation are emitted, but since they do not reach the anode, they must return to the cathode. The space charge (the electrified condition of the space outside the cathode) is negative and thus will repel electrons back to the cathode under any circumstances. Only when the space charge is reduced by the effect of the positive plate voltage does this repelling action relent enough to allow electrons to pass over to the plate.

Dynamic Plate Resistance of the Diode: Small Changes of Current Produced by Small Changes in Voltage.—Often the diode

tube is operated with a combination applied voltage, part of which is steady, the other alternating (as shown in Fig. 56). In this case a current corresponding to each voltage source flows, the steady current I_b caused by the steady voltage E_b and the "signal" current i_p caused by the alternating-current component e_p. These two currents combine in the tube, but usually only the alternating current is of interest, the direct current being used simply to put the tube in operating condition. Hence we devote our attention to the signal voltage e_p and the

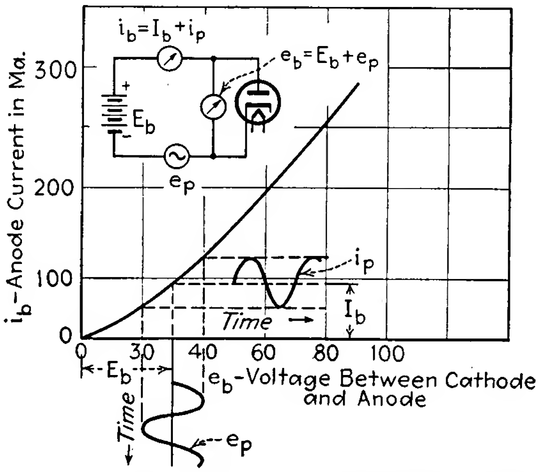


FIG. 56.—Diode action when both direct and alternating voltages are applied simultaneously.

signal current i_p. In particular we consider the change in current di_p caused by a small change de_p in the voltage. The ratio of these small quantities $de_p/di_p = r_p$ is called the *dynamic plate resistance* of the diode. Its value may be found graphically as shown in Fig. 57. In so doing it is found (because of the nonlinearity of the curve) that the value of the ratio is greater when the direct voltage is small than it is when the direct voltage is large. The value of the dynamic plate resistance is thus not constant but varies with the operating condition of the tube, *i.e.*, with the direct-current value of current flowing through it. When the direct-current flow has its rated value, the dynamic plate resistance can be stated as a single value corresponding to this rating.

Use is made of the dynamic plate resistance in conjunction with the external "load" resistance. The sum of these two resistances determines what signal current will flow in response to a definite signal voltage, and this is of course an important consideration in any practical application of the diode. The determination of the direct-current flow also depends on the value of the external resistance, in conjunction with the static resistance of the tube (E_b/I_b). These factors are considered in detail on page 243.

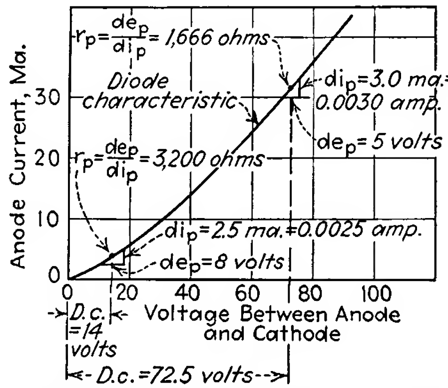


FIG. 57.—Determination of dynamic plate resistance (r_p) from the diode characteristic. As the applied direct voltage is increased, the resistance decreases.

28. The Triode Tube and Its Characteristics.—The triode tube is essentially a diode tube between whose cathode and anode a grid[1] has been inserted. In so far as the cathode and plate are concerned, all of the foregoing discussion of the diode tube applies to the triode tube as well. The current increases with the $\frac{3}{2}$ power of the effective applied voltage, and the dynamic plate resistance of the tube has the same significance both internally and with respect to the external circuit. But the presence of the grid in the triode introduces an extra applied voltage, the grid voltage, and an extra electron-current flow, the grid current. Between the grid voltage, plate voltage, grid current, and plate current there exist three very important relationships and other less important ones. These relationships make the study of the triode much more complicated than that of the diode.

[1] The credit for the invention of the three-element tube is due to Lee de Forest (U. S. Patent 841,387, Jan. 15, 1907).

Van der Bijl's Equivalent Diode. The ³⁄₂-power Law for Triodes.—The ³⁄₂-power law no longer applies in the simple form given in Equation (22), page 58, since the effect of two voltages, those on grid and plate, must be accounted for, whereas (22) contains only one voltage. The extension of the ³⁄₂ power to include the triode was first suggested by Van der Bijl,[1] who adopted the device of considering the triode as an "equivalent diode." In this approach the idea of amplification factor is necessary. The amplification factor expresses the fact (*cf.*

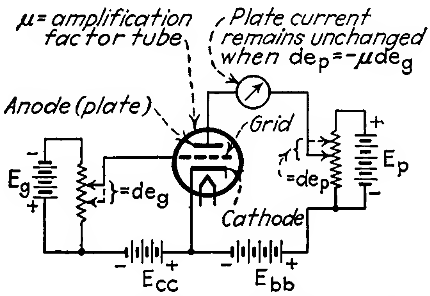


FIG. 58.—Circuit for determining the amplification factor (μ) of a triode. The plate-voltage change, de_p, is found which exactly cancels the effect of a given grid-voltage change, de_g, on the plate current.

page 94) that changes in voltage applied to the grid of a triode are much more effective in changing the plate current than are equal changes in voltage applied to the plate of the tube. The ratio of a small plate-voltage change (producing a given plate-current change) to the grid-voltage change which will produce the same change in the plate current, de_p/de_g, is called the *amplification factor* and is given the symbol μ. If in a given tube a change of 10 volts in the applied plate voltage is required to produce a plate-current change of 1 ma., whereas a change of voltage of 1 volt on the grid would produce the same change in plate current, then the amplification factor of that tube is $\mu = 10/1 = 10$. This quantity is of the greatest value in determining the performance of a tube in a given circuit, but for the present we consider simply its use in developing the ³⁄₂-power law for triodes.

It is found that if a total voltage e_c is applied to the grid of a tube, while a total voltage of e_b is applied to its plate, the effective

[1] VAN DER BIJL, *Verh. d. Phys. Ges.*, **15**, 338 (1913).

voltage e_e tending to draw electrons from the cathode is

$$e_e = \frac{e_b + \mu e_c}{1 + \mu}, \tag{31}$$

where μ is the amplification constant of the structure in question.

Now if this effective voltage is substituted in the $\frac{3}{2}$-power

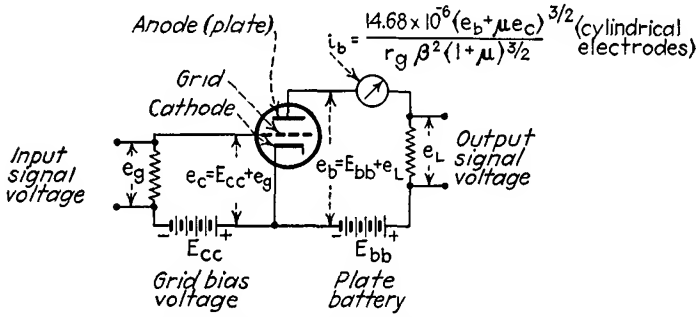


FIG. 59.—Basic connection diagram of the triode tube, showing the dependence of the plate current on grid and plate voltages. For the meanings of symbols, see Appendix II, page 333.

law for diodes [Equations (23) and (24)], we obtain for plane electrodes,

$$i_k = \frac{2.34 \times 10^{-6}(e_b + \mu e_c)^{3/2}}{d_g^2(1 + \mu)^{3/2}} \text{ amp./sq. cm.,} \tag{32}$$

where d_g is the spacing in centimeters between the grid and the cathode. For cylindrical electrodes the equation is

$$i_k = \frac{14.68 \times 10^{-6}(e_b + \mu e_c)^{3/2}}{r_g\beta^2(1 + \mu)^{3/2}} \text{ amp./cm. length.} \tag{33}$$

where r_g is the radius of the grid structure in centimeters. These equations are useful not only in design but for calculating performance. If the value of μ is known, it is possible, for example, to calculate the cathode current i_k leaving the cathode for any desired combination of plate and grid voltage (see Fig. 59), and so set up a group of characteristic curves which describe the operation of the triode over its entire operating range. These curves are discussed in detail on page 119.

It is to be noted that the currents indicated in Equations (32) and (33) are given the symbol i_k which stands for cathode current.

In the diode all the current leaving the cathode reaches the plate, but in the triode the current may split, part of it being collected by the grid, the rest by the plate. In this latter case,

$$i_k = i_c + i_b,$$

where i_c is the total grid current and i_b the total plate current. Normally triodes are operated with a negative voltage applied to their grids, in which case the grid cannot collect a current, $i_c = 0$, and $i_k = i_b$, that is, the total cathode current is identical with the total plate current.

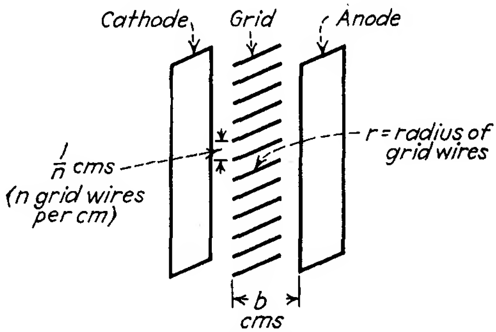


FIG. 60.—Geometrical factors influencing the amplification factor of plane electrodes.

The Determination of Amplification Factor.—The use of Equations (32) and (33) depends on a knowledge of the value of μ, the amplification factor. This value may be measured experimentally, by finding the values of plate- and grid-voltage changes which produce equivalent plate-current changes, as in Fig. 58. This is the most satisfactory procedure, but it is sometimes necessary to calculate the value of μ without experiment. Several expressions have been developed for this purpose, using the dimensions of the grid and plate structures and their positions relative to the cathode as the basis. These expressions are far from satisfactory, since the electrostatic theory on which they rest is highly complicated. The most satisfactory equations are due to Vogdes and Elder[1] and are highly involved hyperbolic functions of the dimensions. The simpler expressions[2] given below are less accurate but much easier to handle with ordinary

[1] VOGDES and ELDER, *Phys. Rev.*, **24**, 683 (1924).
[2] KING, *Phys. Rev.*, **15**, 256 (1920).

mathematics. They give results correct to within 10 or 20 per cent when properly applied:

$$\mu = \frac{0.868\pi nb}{\log_{10}\left(\frac{1}{2\pi rn}\right)}, \tag{34}$$

where n is the number of grid wires per centimeter length of grid, b is the grid-plate spacing in centimeters, and r is the radius of the grid wires in centimeters. The above equation applies to plane electrodes (Fig. 60). For cylindrical electrodes (Fig. 61),

$$\mu = \frac{R_p}{R_g}\left(\frac{0.868\pi nb}{\log_{10}\left(\frac{1}{2\pi rn}\right)}\right), \tag{35}$$

where R_g and R_p are the radii of the plate grid and plate struc-

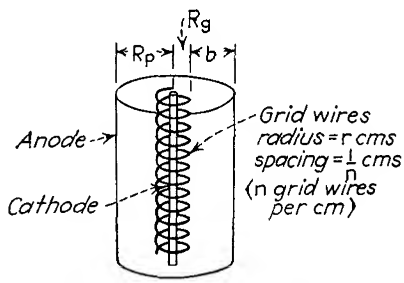


FIG. 61.—Geometry of cylindrical electrodes, used in computing amplification factor.

tures, respectively, measured in the same units. At best these relations are approximate and are used only to establish a working basis for the design.

The Triode Parameters: Dynamic Plate Resistance, Amplification Factor, and Mutual Conductance.—In the preceding discussion it is pointed out that there are three important relationships between the voltages applied to grid and plate and the resulting currents. These three relationships, two of which have already been defined, are the *dynamic plate resistance,* and the *amplification factor,* and the *mutual conductance.* These three *parameters,* as they are called, are usually sufficient to specify the operation of the tube under any given set of conditions.

The dynamic plate resistance of a triode is the same as that of the diode, *i.e.,* it is the ratio of a small change in plate voltage to

the small change in plate current which it produces, $r_p = de_p/di_p$. During the voltage change the grid voltage is maintained constant, since the value of the ratio changes as the total plate current changes, which in turn depends on the applied grid voltage. The dynamic plate resistance is expressed in ohms; its rated values range from below 1000 ohms to above 100,000 ohms depending on the plate-cathode spacing and area. While the operating value of r_p depends, of course, on the grid voltage, the dynamic plate resistance is a parameter which expresses the action of the *plate-current path* of the tube; it is not concerned with the grid voltage except indirectly.

The amplification factor μ of the tube expresses a relationship between a grid-voltage change and the plate-voltage change which has the same effect on the plate current. It thus has to do with the effectiveness of the grid as a control agent, compared with the plate, and it is useful not only in connection with Equations (32) and (33) but also to express the amplification ability of the tube in question. The changes in grid and plate voltages considered are opposite in sign, since the grid must *restore* the change caused by the plate if the total plate current is to remain unchanged during the measurement. Hence $\mu = -de_p/de_g$.

The mutual conductance of a triode is the ratio of its amplification factor to its dynamic plate resistance; it is given the symbol g_m. Actually the mutual conductance is the ratio of a small change in plate current to the small change in grid voltage which produces it, and hence is expressed in microamperes per volt, or micromhos. The ratio expresses the fundamental action of the triode, the control of the plate current by changes in an unrelated voltage applied to its grid. During the changes considered, the total plate current must not change except for the change produced by the grid voltage. Hence the plate voltage must be held constant during the grid-voltage and plate-current changes.

Summary.—Let the symbol d stand, as previously, for "small change in," and let the grid voltage be represented by e_g, the plate voltage by e_p, and the plate current by i_p, then the three triode parameters are defined as:

Dynamic plate resistance,

$$r_p = \frac{de_p}{di_p} \text{ (grid voltage } e_g \text{ constant)} \tag{36}$$

Amplification factor,

$$\mu = -\frac{de_p}{de_g} \text{ (plate current } i_p \text{ constant)} \tag{37}$$

Mutual conductance,

$$g_m = \frac{di_p}{de_g} \text{ (plate voltage } e_p \text{ constant).} \tag{38}$$

Between these three quantities, as can be seen by inspection of the definitions, the following relationships exist:

$$r_p = \frac{\mu}{g_m} \tag{39}$$

$$\mu = r_p g_m \tag{40}$$

$$g_m = \frac{\mu}{r_p} \tag{41}$$

If any two of the three quantities are known, therefore, the other can be calculated.

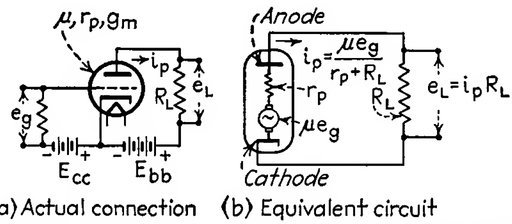


FIG. 62.—Equivalent circuit used in computing the useful amplification of a triode.

The Use of the Triode Parameters. The Equivalent Circuit of the Triode.—To illustrate the use of the triode parameters in predicting the performance of a tube, consider the circuit in Fig. 62*b*. This circuit represents the plate circuit of the tube and the external circuit, *i.e.*, it includes the vacuum path between cathode and plate and the external resistance R_L. Only the alternating "signal" current flowing in this path is considered; its symbol is i_p. Alternating voltage is generated in the circuit by the action of the grid which is represented by the alternating-current generator of voltage μe_g where μ is the amplification factor and e_g is the signal voltage applied to the grid. The resistance of the internal current path in the tube is the dynamic plate resistance of the tube, which is in series with the external load

resistance R_L. The current i_p flowing in the circuit is, by Ohm's law (Chap. XI),

$$i_p = \frac{\mu e_g}{r_p + R_L}. \tag{42}$$

Likewise, the useful voltage appearing across the load resistance is, by Ohm's law,

$$e_L = i_p R_L = \frac{\mu e_g R_L}{(r_p + R_L)}. \tag{43}$$

Finally, the ratio of the useful output signal e_L to the input grid signal e_g across the load resistance (the useful amplification) is

$$\text{Useful amplification} = \frac{e_L}{e_g} = \frac{\mu R_L}{(r_p + R_L)}. \tag{44}$$

The useful amplification of a signal obtained by the use of triode tube is thus expressible in terms of the amplification constant and dynamic plate resistance, taken together with the value of the load resistance into which the tube delivers the amplified signal. Equation (44) is perhaps the most generally useful of all the practical design formulas in electronic-circuit practice. It is used in Chap. XIII.

The Dynamic Grid Resistance of a Triode.—The foregoing discussion has neglected grid current, since ordinarily no appreciable grid current flows in amplifier circuits. In certain high-efficiency amplifiers (class *B* and *C* amplifiers, see page 277), in grid-leak detectors, and in oscillators, the applied alternating voltage is large enough so that the grid becomes positive, with respect to the cathode, for a part of each signal cycle. When the grid is positive it collects a part of the cathode-emission current, thus reducing the current available to the plate. More important, the current flowing through the grid-cathode space must be supplied by the signal circuit feeding the grid of the tube, and the design of this circuit must take this fact into account. In this problem the dynamic grid resistance of the tube is important; it is the ratio of a small grid-voltage change to resulting change of current flowing to the grid:

$$r_g = \frac{de_g}{di_g}. \tag{45}$$

Ordinarily this factor may be neglected; when it is taken into account, it may be added in series with the alternating-current resistance of the input circuit which applies voltage between the grid and cathode.

A Problem in Tube Design and Application.—The following rather comprehensive problem is not illustrative of present electronic engineering practice, but it does illustrate the applications of the principles stated in the foregoing sections.

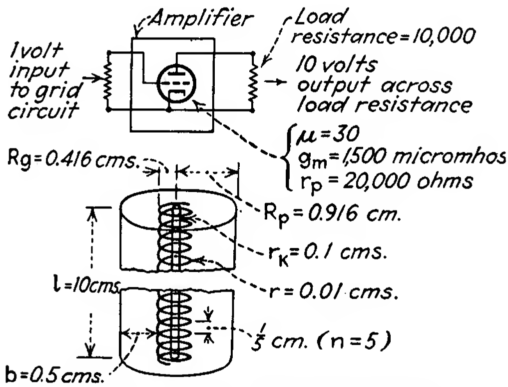


FIG. 63.—Triode tube design (Prob. 2).

Problem 2. (Refer to Fig. 63.) It is required to build an amplifier circuit which will work into a 10,000-ohm load resistance and which will have a useful amplification of 10 times. Determine the amplification factor, plate resistance, and mutual conductance of the required tube. Suggest a tube design (dimensions and spacings of elements) which would provide these characteristics.

Given:

$$\frac{e_L}{e_g} = 10.$$

$$R_L = 10{,}000 \text{ ohms.}$$

To find:

μ, r_p, g_m, and tube dimensions.

Rewrite Equation (44) thus:

$$\frac{e_g}{e_L} = \frac{r_p + R_L}{\mu R_L} = \frac{1}{g_m R_L} + \frac{1}{\mu}. \tag{44a}$$

$$\frac{1}{10} = \frac{1}{10{,}000 g_m} + \frac{1}{\mu}.$$

This indicates that g_m and μ can be specified independently under the given conditions. Experience indicates, however, that a g_m value

higher than 2000 micromhos is difficult to achieve. Assume a value of g_m of 1500 micromhos (0.0015 mho). Then

$$0.10 = \frac{1}{15} + \frac{1}{\mu} = 0.066 + \frac{1}{\mu}$$

$$\frac{1}{\mu} = 0.033; \qquad \mu = 30.$$

Then

$$r_p = \frac{\mu}{g_m} = \frac{30}{0.0015} = 20{,}000 \text{ ohms.}$$

Therefore,

$$\mu = 30; \qquad r_p = 20{,}000 \text{ ohms}; \qquad g_m = 1500 \text{ micromhos.}$$

To find:

A set of tube dimensions satisfying these values.

The equation for μ, assuming a cylindrical structure as a convenient form, is

$$\mu = \frac{R_p(0.868\pi nb)}{R_g \log_{10}[1/(2\pi rn)]} \tag{35}$$

This equation admits of a wide choice of values. Assume that the grid-wire radius r is 0.01 cm., that the number of grid wires is 5 per centimeter ($n = 5$), and that the grid-plate spacing b is 0.5 cm.

Then, substituting these values in (35), we obtain a value of the ratio R_p/R_g as follows:

$$30 = \frac{R_p}{R_g}\frac{(0.868 \times 3.14 \times 5.0 \times 0.5)}{\log_{10}[1/(2 \times 3.14 \times 0.01 \times 5.0)]}$$

$$= \frac{R_p}{R_g}\frac{6.8}{\log_{10} 3.18} = \frac{R_p}{R_g}\frac{6.8}{0.5}.$$

$$\frac{R_p}{R_g} = 30\,\frac{0.5}{6.8} = 2.2 \text{ times.}$$

Now the difference between R_p and R_g is b. Hence

$$R_p - R_g = 0.5.$$

But

$$2.2R_g - R_g = 0.5.$$

$$R_g = \frac{0.5}{1.2} = 0.416 \text{ cm.}$$

$$R_p = 2.2R_g = 0.916 \text{ cm.}$$

The dimensions of the system have now been fixed, except the length of the elements, l. This factor is decided by the necessity of obtaining values of $r_p = 20{,}000$ ohms, and $g_m = 0.0015$ mho.

First calculate the cathode current, i_k, per centimeter length of element, by Equation (33):

$$i_k = \frac{14.68 \times 10^{-6}(e_b + \mu e_c)^{3/2}}{R_g\beta^2(1 + \mu)^{3/2}} \text{ amp./cm. length.} \tag{33}$$

Assuming that the cathode is 0.1 cm. in radius, then the ratio R_p/R_k is 9.16 and the value of β^2 (taken from Fig. 30) is approximately 0.955. R_g is 0.416 cm. and the amplification factor is 30. Hence

$$\begin{aligned} i_k &= \frac{14.68 \times 10^{-6}}{0.416 \times 0.955 \times (31)^{3/2}} \times (e_b + 30e_c)^{3/2} \\ &= 2.14 \times 10^{-7}\ (e_b + 30e_c)^{3/2} \end{aligned}$$

Values of e_b and e_c must now be inserted in this expression. It will be remembered, since $r_p = de_p/di_p = 20{,}000$ ohms, that a change of 1 volt in e_b must cause a change of 1/20,000 amp. in the total cathode current $i_k \times l$, where l is the length of the elements, provided that the cathode current all goes to the plate (*i.e.*, assuming a negative grid). Assume a total grid voltage e_c of -2.0 volts. Then if the plate voltage e_p is 200 volts, the cathode current per centimeter length is

$$\begin{aligned} i_k &= 2.14 \times 10^{-7}(200 - 60)^{3/2} \\ &= 3.54 \times 10^{-4} \text{ amp./cm. length.} \end{aligned}$$

A change of 1 volt in e_b, to 201 volts, increases the current to

$$\begin{aligned} i_k &= 2.14 \times 10^{-7}(201 - 60)^{3/2} \\ &= 3.60 \times 10^{-4} \text{ amp./cm. length.} \end{aligned}$$

The change in i_k is 0.06×10^{-4} amp. based on a structure 1 cm. long. The necessary change is $1/20{,}000 = 0.5 \times 10^{-4}$ amp. The required current change is thus $0.5/0.06 = 8.3$ times as great.

Now, g_m the mutual conductance, must have a value of 0.0015 mho, *i.e.*, a 1-volt change in e_c must cause a cathode-current (plate-current) change of 0.0015 amp. Changing the grid voltage to be -0.1 volt, the cathode (plate) current becomes

$$\begin{aligned} i_k &= 2.14 \times 10^{-7}(200 - 30)^{3/2} \\ &= 4.71 \times 10^{-4} \text{ amp./cm. length.} \end{aligned}$$

The change in current is

$$(4.71 - 3.54) \times 10^{-4} = 1.17 \times 10^{-4} = 0.000117 \text{ amp.}$$

The required current change is $0.0015/0.000117 = 12.8$ times as great as against 8.3, the value arrived at by considering the dynamic plate resistance. The average of these values is about 10 times.

To obtain this 10 times increase in current change, the length l of the element structure must be 10 cm., an impractically large value. To avoid this difficulty, another design must be attempted, using a smaller value of b, the plate-grid spacing, which will lead to a more compact structure, and hence a larger current per centimeter length of element.

The above example, it should be remembered, is simply a theoretical illustration of the use of the triode equations. In the design of practical tubes, much use is made of experience, and mathematics, where used, is on a more complex plane than that given above. For the purposes of this discussion, the example serves simply to show how the various factors in the tube design influence its characteristics.

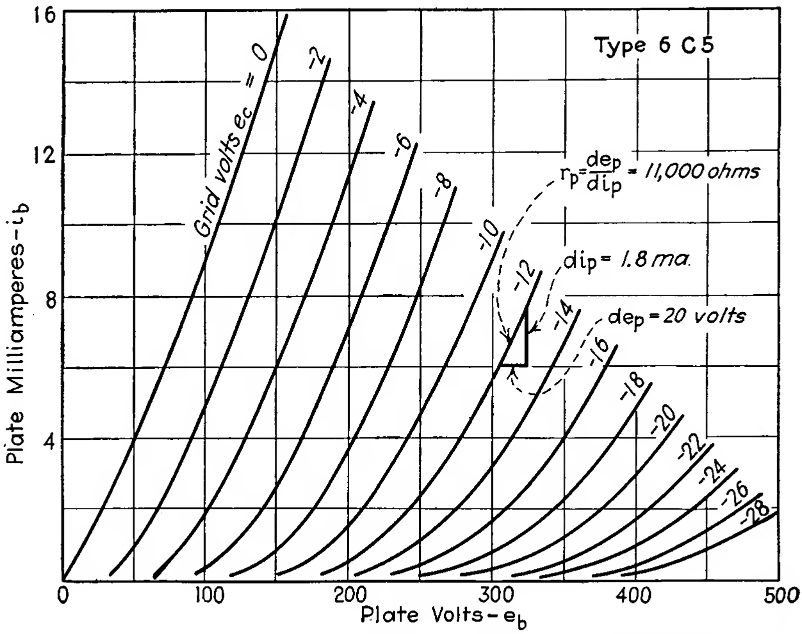


FIG. 64.—Triode curves, the i_b-e_b family. The dynamic plate resistance may be computed from the curves as shown.

The Characteristic Curves of the Triode.—The triode parameters μ, g_m, and r_p are not constant in any tube, but depend on the operating conditions, *i.e.*, upon the voltages applied to the grid and plate. Ordinarily these applied voltages are the recommended operating values, and it is then possible to use the corresponding parameter values, as stated by the manufacturer of the tube.

Whenever other than rated voltages are applied, and whenever the tube operates over wide ranges of values (*i.e.*, whenever the changes in signal voltage are comparable in size with the steady

direct voltages applied to the tube) then the parameter values stated as "rated" no longer apply. To determine the operation of the tube in these cases, resort is taken to a set of characteristic curves, which show the various currents and voltages in their proper relationship over the entire range of possible values.

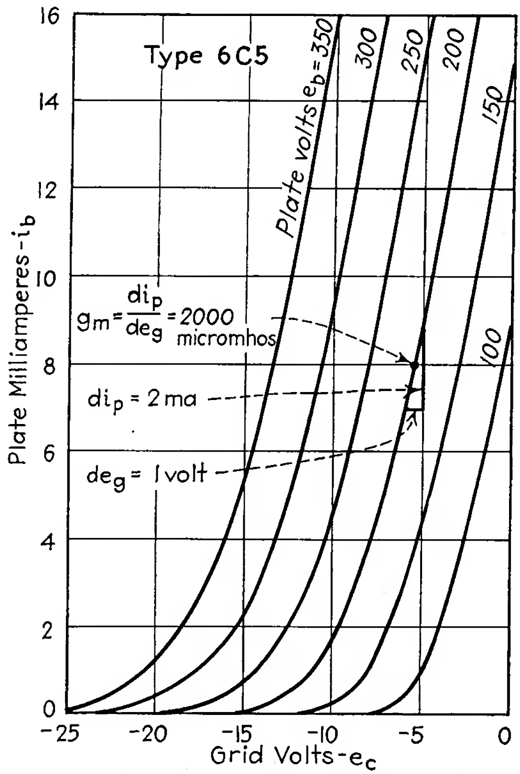


FIG. 65.—Triode i_b-e_c curves. The mutual conductance at 8 ma. and 200 volts is determined as shown.

Since a family of plane curves is capable of representing three quantities simultaneously, it is possible to select any combination of three currents and voltages. The voltages selected are the total grid voltage e_c, the total plate voltage e_b, and the total plate current i_b. Grid current i_c is seldom considered but may be treated similarly in combination with the total grid voltage and the total plate voltage.

The graphical relationship between i_b, e_b, and e_c can be represented in three ways. The first (Fig. 64) is a series of i_b-e_b curves for values of e_c, and is called the plate family. The inverse slope of these curves at any point gives the dynamic plate resistance r_p at that point. The second family (Fig. 65) is composed of i_b-e_c curves for different values of e_b and is

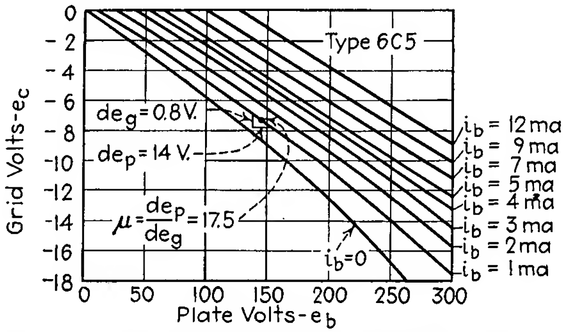


FIG. 66.—Triode e_b-e_c curves, from which the amplification factor may be found.

called the grid-voltage plate-current family. The slope of any of these curves at any point is the value of the mutual conductance g_m at that point. The third (Fig. 66) is a family of

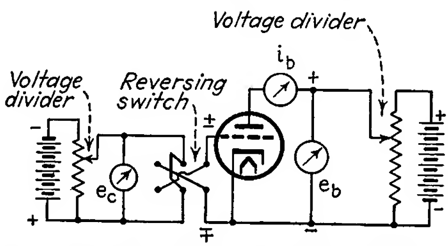


FIG. 67.—Circuit for measuring points on triode curves.

e_b-e_c curves for different values of i_b and is called the grid-plate-voltage family. The inverse slope of any curve in this set is the amplification factor μ at that point.

The shape of these curves may be explained in terms of Equations (32) and (33), but ordinarily they are experimental curves derived from measurements. The circuit shown (Fig. 67) is used to perform these measurements. The measured values are representative of the tube alone and independent of the manner in which it is loaded. For this reason the curves

are called *static characteristics.* To obtain the performance of the tube in a definite circuit, of course, it is necessary to consider the static curves in relation to the load resistance, and in this case a different set of curves, called *dynamic char-*

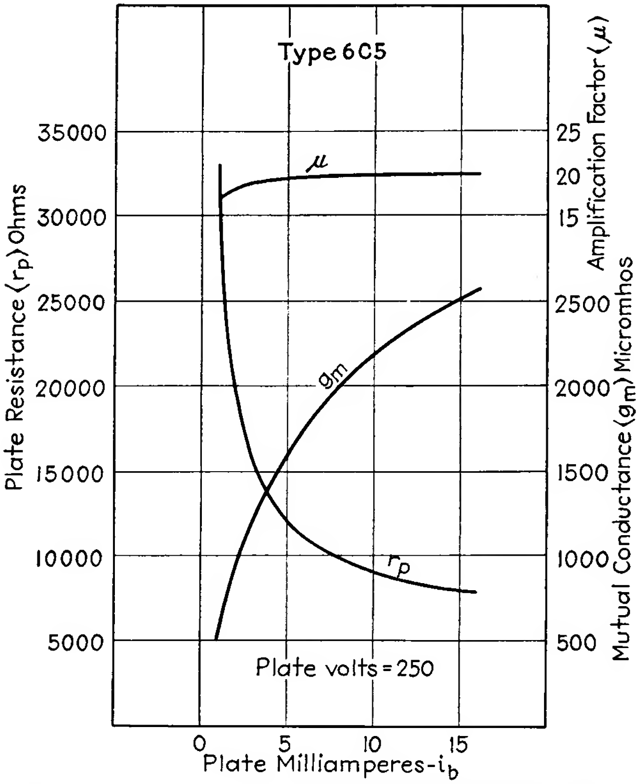


FIG. 68.—Variation of triode parameters with plate current.

acteristics, results. Since the dynamic characteristic of a triode is, strictly speaking, a matter of its application, the discussion of it is deferred to Chap. XI.

A highly useful method of indicating changes in μ, r_p, and g_m is given in Fig. 68, in a series of curves plotted against plate current.

29. The Tetrode Tube and Its Characteristics.—In Chap. XI, the action of a capacitor, (two conducting plates separated by an insulator) is discussed, and its ability to "pass" alternating

current is explained in terms of the alternate charging of the plates. The elements in a vacuum tube are actually capacitor plates, and as a consequence there is a tendency for alternating current to pass between the grid and the plate, entirely regardless of the electronic action in the vacuum space. When this grid-plate alternating current flows, the independence of the grid circuit from the plate circuit is destroyed. When a high degree of amplification is desired, the large voltages in the plate circuit will, therefore, be partially imposed on the relatively weak grid voltage, through the grid-plate capacitance. This not only destroys the faithfulness of the amplification but also, if allowed to progress far enough, will set up a *self-sustained* signal. The plate-voltage changes thus act through the grid-plate capacitance to set up a fictitious signal in the grid circuit, which has no relationship to the desired signal. This condition, called oscillation, has its uses in oscillator circuits, but in amplification circuits it is a very important limitation to the upper limit of amplification obtainable.

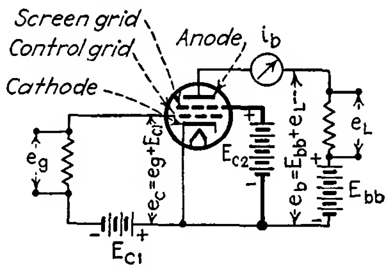


FIG. 69.—Basic connection of the tetrode tube. The heavy line indicates the sole difference between the triode and tetrode connections.

By placing another grid between the control grid and the plate and by maintaining it free of any alternating current, it is possible to reduce the capacitance which exists between control grid and plate. Since this extra grid serves the purpose of shielding the plate from the control grid, it is called a "screen" grid. Usually this grid is given a positive voltage, with respect to the cathode (see Fig. 69), so that it will accelerate the electrons toward itself. The electrons are collected by the screen grid only in part, however. Most of them pass through the spaces in the grid and are collected by the plate beyond. The reduction of grid-plate capacitance so obtained makes the grid-to-plate reaction small and permits very large values of amplification to be obtained.

The screen grid has another effect of the greatest importance. By shielding the plate from the cathode, it makes the plate voltage much less effective in attracting electrons from the

cathode, whereas the ability of the control grid is not impaired. Therefore the grid is a much more effective agent in changing the plate current than is the plate, and as a result the amplification constant of the tetrode tube is high. The tube thus inherently has good amplifying properties, the use of which it permits because of its low plate-to-control-grid capacitance. The characteristics of the tetrode are high amplification factor, high dynamic plate resistance, and average values of mutual conductance; it is especially valuable as an amplifier of voltage and power.

One limitation of the tetrode is the generation of secondary electrons, which are emitted from the plate of every tube when the main electron stream bombards it. In the triode these secondary electrons are immediately recollected by the plate, since the plate is the only element in the tube at a positive potential with respect to the cathode. In the tetrode, however, the presence of the positively charged screen grid near the plate permits the collection of some of the secondary electrons by the screen grid. This current of electrons from the plate to the screen is in effect a "negative" plate current, *i.e.*, it subtracts from the effective current entering the plate. This undesirable effect is minimized in the pentode tube, discussed in the next section.

The Characteristic Curves of the Tetrode.—The addition of the extra grid makes it necessary to consider a new voltage (the screen voltage) and a new current (the screen current) in addition to the currents and voltages present in the triode. With three voltages and three currents to consider, the complete characteristic curves of a tetrode must be expressed in three distinct groups, each group containing several families of curves. Fortunately in ordinary cases, control-grid current is not present, the screen voltage has a fixed value, and the screen current is of no interest (at least so far as the signal components are concerned). Hence we find that the factors to be considered are the plate and control-grid voltages (e_b and e_c) and the current is the plate current (i_b). One family suffices therefore to express the relationship between these three factors. The curves are exactly the same as those of the triode, except for differences in shape. In other words, for all practical purposes the tetrode may be considered as an "equivalent" triode, and the relation-

ships given in the section on triodes then apply directly to the operation of the tetrode. The only change is the difference in the shape of the characteristic curves, due to the effect of the screen grid in shielding the plate and in collecting a current which subtracts from the plate current.

The plate-family curves of a typical tetrode are shown in Fig. 70. Compared with the curves of the triode, the chief difference is their flatness (corresponding to a high dynamic plate resistance, caused by the shielding effect of the screen

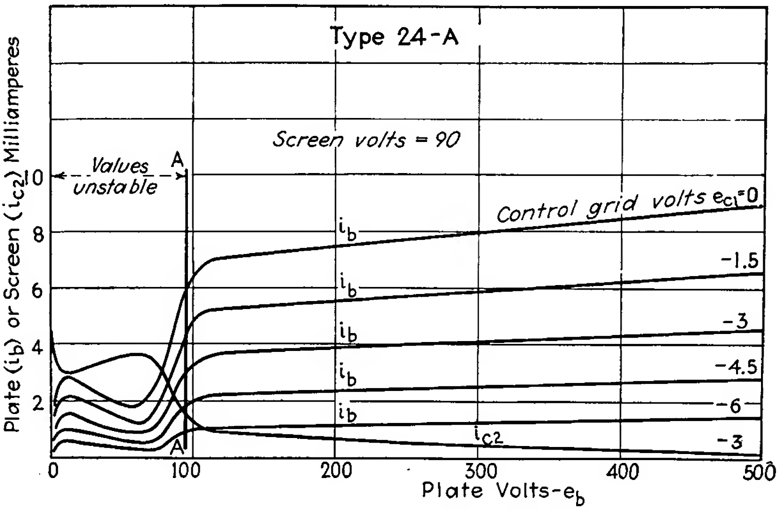


FIG. 70.—Plate family of a typical tetrode. One screen-grid current curve is included.

grid) and the dip in the plate current at low plate voltages (due to the collection of secondary electrons by the screen grid). The values of μ, g_m, r_p, have the same significance as in the triode, and are applied in the computation of useful amplification (page 115) in the same way.

30. The Pentode Tube and Its Characteristics.—The suppression of the secondary electron current, which flows from the plate to the screen-grid of a tetrode, is accomplished by placing a third grid, called the suppressor grid, between the plate and the screen grid and giving this grid a voltage considerably lower than that applied to the plate. Usually the suppressor grid is connected directly to the cathode (see Fig. 71), *i.e.*, its potential with respect to the cathode is zero. Since there is no tendency

of this grid to collect electrons, the secondary electrons return to the plate, for the most part, although some of them may go through the suppressor grid and be collected by the screen grid. The pentode tube has all the special characteristics of the tetrode, therefore, except that the secondary electron current flowing from the plate is reduced to a very small amount. Consequently the characteristic curves of the pentode (Fig. 72) are very similar to those of the tetrode, except that the dip in the plate current at low plate voltages does not occur.

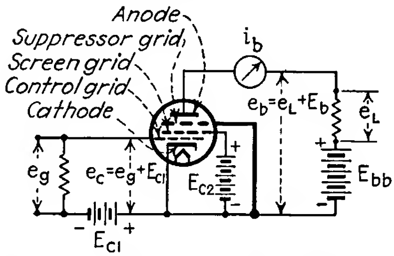


FIG. 71.—Basic connection of the pentode tube.

The presence of the suppressor grid still further increases the shielding between the plate and the cathode, and by the same

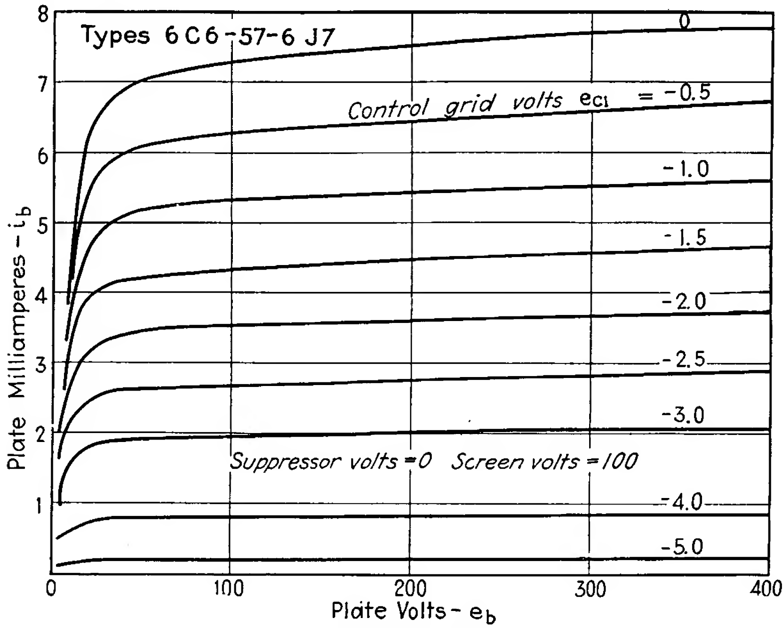


FIG. 72.—Plate family of a typical pentode tube

reasoning as in the case of the tetrode, the amplification factor and the dynamic plate resistance are thereby both increased. In general, therefore, the pentode tube is distinguished by very

high amplification factor (up to 1500 in some types) and high plate resistance, with average values of mutual conductance.

Since the suppressor-grid voltage and the suppressor-grid current are ordinarily of no interest, in connection with the signal components in the tube at least, the pentode can be described in terms of the triode parameters μ, r_p, and g_m, and these quantities have the same significance in circuit analysis as they do in the triode. The triode tube is, in fact, the prototype of all amplifying tubes, be they triodes, tetrodes, or pentodes.

31. Practical Constructions in Triodes, Tetrodes, Pentodes. The Remote Cutoff, Supercontrol, and Beam-power Structures.—The construction of triode-, tetrode-, and pentode-element structures varies widely with the intended use of the tube. In general there are two types of tube: voltage amplifiers and power amplifiers. The voltage amplifiers are used primarily for producing a high amplification of the input voltage, hence the output signal current need not be large. Consequently the dynamic plate resistance of these tubes may be large, and the emission current from the cathode comparatively small. Generally a high amplification factor is desired.

The power amplifiers, on the other hand, are used for converting the input signal voltage into an output signal having both large voltage and large current values, since the product of the current and the voltage represents the power output. These tubes have cathodes capable of supplying high emission (see Fig. 73) and plate structures capable of withstanding the heavy electron bombardment which accompanies the current flow. Usually the plate is treated with a carbon layer, to render it a better dissipator of heat and also to reduce the tendency to liberate secondary electrons. The grid structures must also be properly designed so that they do not assume high temperature during operation, else they will begin to emit electrons thermionically, and grid current will flow between grid and plate.

Pains must also be taken to secure and maintain the proper vacuum within the tube. The technique of glass working in glass tubes and of welding in metal tubes is highly developed in modern tube manufacture. Included in the tube is the "getter," a magnesium layer sputtered on the inside of the tube to absorb gases which may remain after the exhaust process or which may be liberated from the electrodes during operation. To guard

against the latter contingency, the elements are treated with a hydrogen-firing process before assembly within the tube, and with an induction-heating process during the exhaust of the tube. The details of the tube-manufacturing process are of great interest but unfortunately cannot be described here because of the greater importance of material having to do with tube applications.

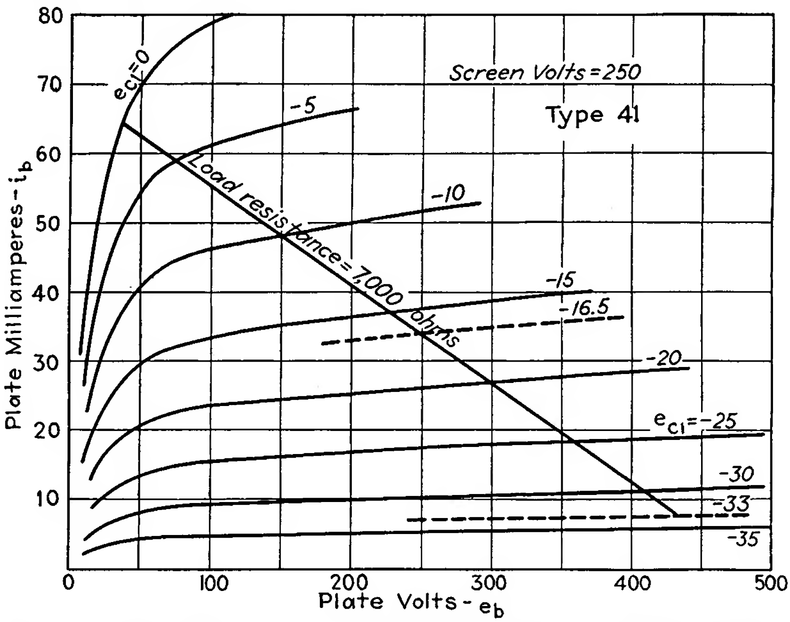


FIG. 73.—Plate family of a typical power amplifier pentode. Compare the values of plate current with those in Fig. 72.

Figure 74 has been prepared to show the μ, r_p, and g_m values of some 50 of the more commonly used low-power thermionic vacuum tubes, and to point out the difference in these values between triodes, tetrodes, and pentodes, and between voltage and power amplifiers. This figure will repay considerable study.

Remote Cutoff and Supercontrol Construction.—The point in the characteristic curve at which the plate current becomes zero is called the cutoff point, and the value of grid voltage (at some specified plate voltage) at which the cutoff occurs is called the cutoff grid voltage. By proper design of the control grid it is

possible to place the cutoff point at almost any desired value of grid voltage. In particular the sharp-cutoff designs are arranged so that the plate current reduces to zero sharply as the grid voltage is made more negative. In contrast, the remote

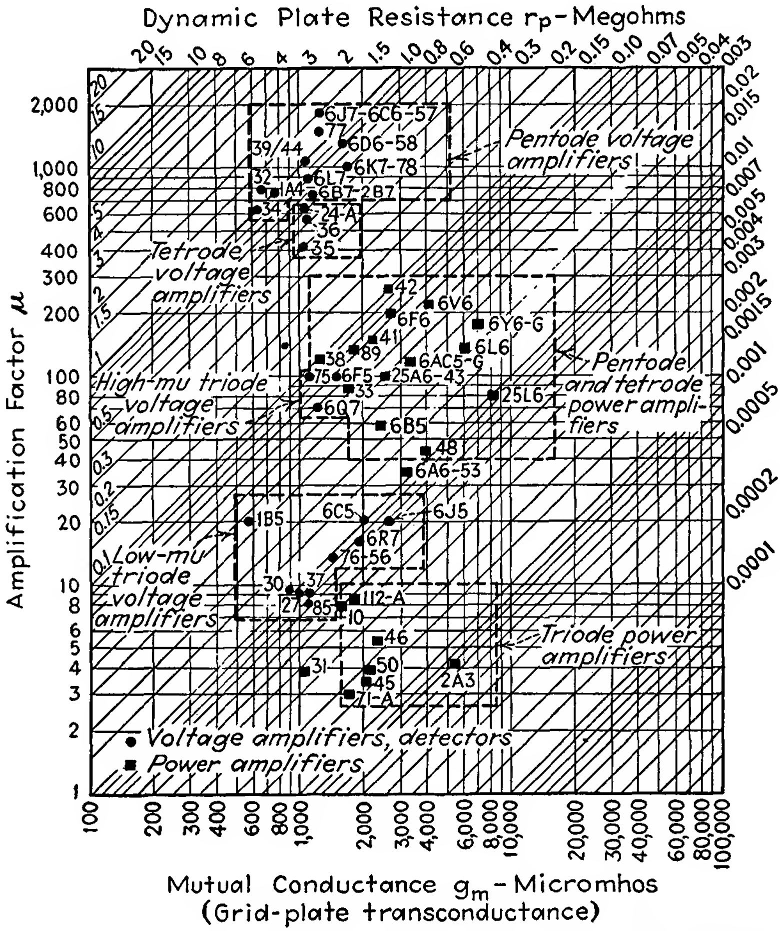


FIG. 74.—Representative values of amplification factor, dynamic plate resistance, and mutual conductance of receiving tubes.

cutoff types are designed so that the current does no cutoff until extremely large grid voltages are applied, as shown in Fig. 75.

The remote-cutoff tubes are particularly suitable for what is known as *supercontrol* or *variable*-mu design. By winding the grid

helix with an increased winding pitch in the center, the effectiveness of the grid in controlling the current can be made to vary by changing the d-c value of grid voltage. This is in effect a change in the amplification factor of the tube (hence the name "variable mu"), under the control of an external circuit. Many uses of this effect have been found; the most important

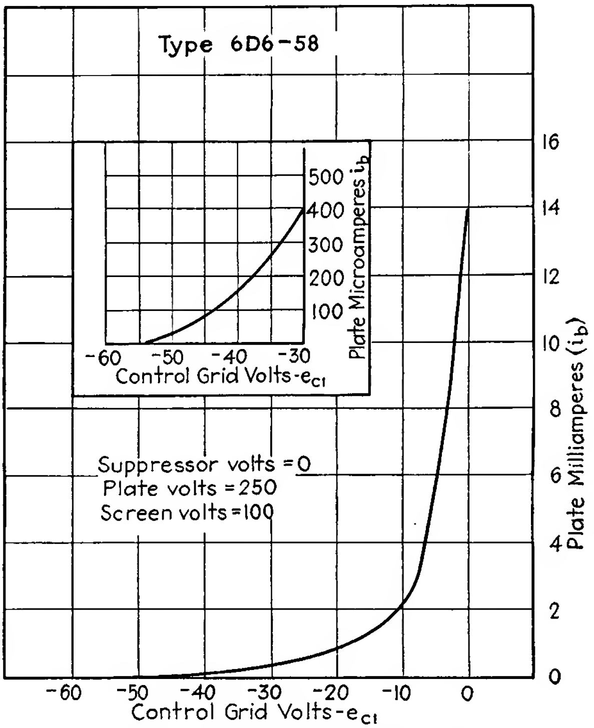


FIG. 75.—Remote cut-off characteristic.

application is the automatic volume control (a-v-c) used in all modern radio receivers, which adjusts the amplification of the signal in accordance with its strength as it comes from the antenna. Details of this arrangement are given in Chap. XIII.

The Beam-power-tube Construction.—The application of electron-beam-producing principles to conventional tubes has resulted in the development of the beam-power tube, a tetrode which operates like a pentode and which has advantages over the ordinary pentode. Specially designed cathode, control-grid,

and screen-grid structures produce beams of electrons which result in the production of regions of high electron density between the plate and screen grid (see Fig. 76). Since high electron density corresponds to a low potential (high negative space charge), there exists between the plate and screen a region of low potential which acts on the secondary electrons in exactly the same way as a suppressor grid. However, since no actual suppressor-grid structure is present, the effective area of the plate for collecting electrons is much larger than in a conventional pentode, and furthermore, the suppressor action is much more complete because the suppressing field is continuous rather than distributed by wires as in the case of the conventional grid structure. These advantages have been combined with a large and efficient cathode structure, and with precautions to insure low operating temperatures of the grids and the plate. The result is a tube which has very great power-handling capacity in relation to its size. Its *power sensitivity* (ratio of power output to the grid-signal voltage) is higher than that of any other receiving-type tube.

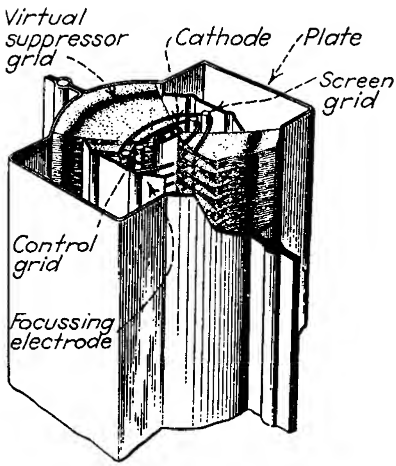


FIG. 76.—Beam-power element structure.

Other Constructions.—Multifunction tubes containing one or more distinct sets of elements within a single envelope have already been mentioned (page 95). One specialized form of double tube, the coupled triode, contains two triodes within a single envelope, the first directly coupled to the second, and supplying both direct- and alternating-current components to the grid of the second tube.

Problems

1. Determine the emission efficiency of an indirectly heated cathode:

Area: 0.5 sq. cm.
Coating: Ba-Sr oxide ($A = 10^{-2}, b_0 = 12{,}000$)

Temperature: 1100°K.
Resistance of heater wire: 21 ohms.
Heater current: 0.3 amp.

2. Compute and plot the i_b-e_b curve of the following diode tube:

Cylindrical anode: radius 1.0 cm., length 2.0 cm.
Cylindrical cathode: radius 0.1 cm., length 2.0 cm.
Consider values of e_b from 0 to 200 volts.
State the saturation current if the voltage at which saturation occurs is 560 volts.

3. From the curve computed in Prob. 2, compute the average output current when the applied voltage is 110 volts, 60 cycle (peak value 154 volts). (HINT: Using paper ruled in fine squares, draw the current pulses produced by the applied alternate current, count the squares under the curve of one pulse, divide by the number of squares along the axis between the beginning of one pulse and the beginning of the next.)

4. Compute the dynamic plate resistance of the diode in Prob. 2 at a voltage of 100 volts.

5. Compare the useful amplification obtained from:

A type-6C5 triode,
A type-24A tetrode,
A type-6D6 pentode,

when operated at rated voltages into a 100,000-ohm load. Compute the values of load resistance required with each of these tubes to produce a useful amplification equal to 50 per cent of the amplification factor of the tube. (Obtain tube constants from Fig. 74, page 129.)

6. Repeat Prob. 2 (page 116) for the following values: Required load resistance, 25,000 ohms; amplification, 15 times. Assume the following values: $g_m = 1500$ micromhos; $r = 0.01$ cm.; $n = 5.0$; $b = 0.3$ cm.; $e_b = 200$ volts; $e_c = -1.0$ volt; $\beta^2 = 0.7$.

Find μ, r_p, R_g, R_p, and l.

7. Determine the values of g_m, r_p, and μ from Figs. 64, 65, and 66 at $e_b = 150$ volts, $i_b = 6$ ma., $e_c = -4$ volts, by graphical construction.

Bibliography

VAN DER BIJL, H. J.: "Thermionic Vacuum Tube," McGraw-Hill Book Company, Inc., New York, 1920.

CHAFFEE, E. L.: "Theory of Thermionic Vacuum Tubes," McGraw-Hill Book Company, Inc., New York, 1933.

EASTMAN, J. V.: "Fundamentals of Vacuum Tubes," McGraw-Hill Book Company, Inc., New York, 1937.

MACARTHUR, E. D.: "Electronics and Electron Tubes," John Wiley & Sons, Inc., New York, 1936, Chaps. IV, V.

DOW, W. G.: "Fundamentals of Engineering Electronics," John Wiley & Sons, Inc., New York, 1937, Chaps. II, VI, VII.

CHAPTER VII

GAS-FILLED THERMIONIC ELECTRON TUBES

Introduction.—Gas-filled tubes are similar to the thermionic vacuum tubes in that both contain a cathode and an anode as the basic elements, and both may contain one or more grid structures for control and regulatory purposes. But here the resemblance ends. Vacuum- and gas-filled tubes are quite different in their operating properties and in their applications. The triode vacuum tube, we remember from the preceding chapter, is capable of controlling small amounts of power with a very fine degree of detail. Throughout the operating range of such tubes the control is smooth and continuous. These tubes are fitted for the purposes of communication engineering, since communication currents are of great complexity and, when amplified, must be reproduced accurately. The amount of power transfer is small and low efficiency in the transfer can be tolerated. For industrial purposes, on the other hand, large power-handling capacity and high efficiency of operation are both of great importance. For such purposes, the high-vacuum tube is not suited, except for specialized control of small amounts of power.

Gas-filled tubes, in contrast, have limited application in communications practice, but they are eminently suited to the heavy-duty service required in industrial applications. Gas-filled tubes are suited to these larger jobs for one principal reason: The voltage drop between cathode and anode, when they are conducting current, is much lower than it it is in a vacuum tube. This low voltage drop represents a comparatively small power loss within the tube, and the efficiency of power transfer is correspondingly high. For this reason, gas-filled tubes can be operated with much higher cathode-to-anode currents than could a vacuum tube of the same size. Very large currents can be carried in vacuum tubes, of course, but only at such high voltages, and with such large and costly element structures, that the cost (both first cost and operating cost) is prohibitive. Hence

gas-filled tubes are used, for reasons of economy, in heavy-current applications. But the gas-filled tubes have one very important disadvantage: The control of the power flowing through the tube is not smooth and continuous. In fact, the control is so rough and imperfect that, in general, gas-filled tubes cannot be used for amplification and hence are of little value in communication circuits. The function of the gas-filled tube is limited to turning current on and off, in much the same manner as in an ordinary electromechanical relay. Gas-filled tubes are much more versatile than the simple relay, of course, but they cannot approach the triode vacuum tube in its ability to increase or decrease the current flow by any amount within the operating range. Fortunately in most industrial applications this fine degree of control is not required, so that the gas-filled tube's usefulness is not curtailed.

Types of Gas-filled Tubes.—The members of the gas-filled group of tubes are classified according to the type of electron emission employed and according to the number of elements present. The *cold-cathode* types, of relatively minor practical importance, obtain emission from a cold surface by field emission and by bombardment of the positive ions. Usually the cold-cathode tubes have but two elements, the anode and cathode, but certain types contain a control element and so can act as relays.

The most important tubes in the group are the *gas-filled thermionic tubes,* which contain hot cathodes. The cathode structures are usually of the oxide-coated type, varying in form from the simple coated ribbon to highly complicated heat-conserving structures. The gas-filled thermionic tubes are built as diodes, as triodes, and as tetrodes. The diode types are used simply for rectification, while the triodes and tetrodes are used for controlled rectification and in a variety of relay and power-transformation circuits.

The third class within the gas-filled group is the *mercury-pool tube.* These tubes obtain electron emission from a pool of liquid mercury. The emission from a single cathode spot (of the field-emission type, see page 49) is limited to about 15 amp.; if more current is required, two or more spots will form until the current requirement of the external circuit is satisfied. The mercury pool is thus a very rugged and flexible form of high-

current cathode. The pool-type tubes are accordingly used in heavy-current service.

The accepted name of the diode form is *phanotron*, of the triode and tetrode forms of thermionic gas-filled tubes it is *thyratron*, while that of the triode type of mercury-pool tube is *ignitron*.

In addition to these forms of electron tubes used for power control and transformation, there is a large family of gas-filled lamps, used primarily for the production of light, which are considered separately in Chap. IX, and a group of gas-filled phototubes, described in Chap. VIII.

32. Gas-filled Thermionic Diodes.—The gas-filled thermionic diode (phanotron) contains a thermionic cathode and an anode, sealed within a glass or metal container and immersed in a gas or vapor at low pressure. Mercury vapor is most commonly used, the vapor being obtained from a small amount of liquid mercury sealed within the tube. The pressure is the vapor pressure of the liquid, and hence depends upon the temperature, varying from 0.0002 mm. (0°C.) to 0.1 mm. (85°C.). To avoid the resulting temperature coefficient, several forms of inert gas are used in place of mercury vapor, neon, argon, and helium being the most common. The pressure of these gases is low, well below that of the minimum sparking potential of the gas (see page 87), usually in the range from 0.15 to 0.5 mm.

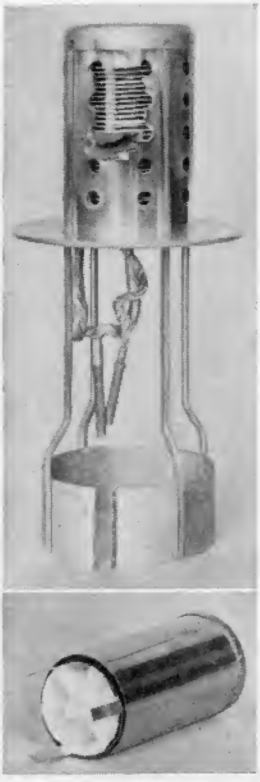


FIG. 77.—High-current cathodes used in gas-filled tubes.

Two typical cathode structures used in gas-filled tubes are shown in Fig. 77. They are designed to produce large amounts of emission and hence have large surface areas. The current-carrying capacity of the tube is limited by the emitting ability of the cathode, and it is important that all the required current be available thermionically. If other forms of emission appear, they will damage the cathode and hence must be avoided.

The anode of the tube, made of metal or graphite, must be large enough to operate without undue heating, since it is important that the anode remain inactive except in conducting the electron current to the external circuit. Any tendency of the anode to emit electrons reduces the tube's usefulness as a rectifier.

The operating voltage between the cathode and anode during current conduction depends upon the gas or vapor used. In mercury tubes the range is from 6 to 20 volts, depending on the temperature (Fig. 78); in neon and helium tubes the voltage

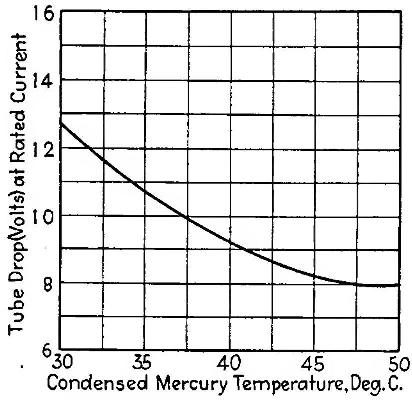


FIG. 78.—Voltage drop during conduction, *vs.* temperature of condensed mercury, at rated current. (Type FG-42.)

drop varies from 20 to 25 volts. The current capacity varies with the size of the tube: a typical 4-in. tube will handle about 0.5 amp., while a 25-in. tube can carry 75 amp., average current. Much higher currents, up to 500 amp., may be carried for very short periods.

The ability of gas-filled diodes to carry large currents is, of course, no recommendation in itself. Copper wire will do the same thing, and with greater efficiency. But gas-filled diode tubes have one special ability which wire does not possess; the ability to carry current in one direction only. This implies that the cathode of the tube is a good emitter of electrons and that the anode is a very poor emitter. Under such circumstances current will flow only when the anode is more positive than the cathode. Hence, when alternating voltage is applied between the electrodes, the current will flow only during the half cycles

which make the anode positive. This is the fundamental rectifier action.

Ordinarily no current flows when the anode is negative. However, if a very large negative voltage is placed on the anode, current conduction may start even if the anode is a poor thermionic emitter. This follows from the fact that a current will be carried between any two electrodes immersed in a gas, whether they be good emitters or not, if a sufficiently high voltage is applied between them. In practice this fact limits the maximum negative voltage which may be placed on the anode to about

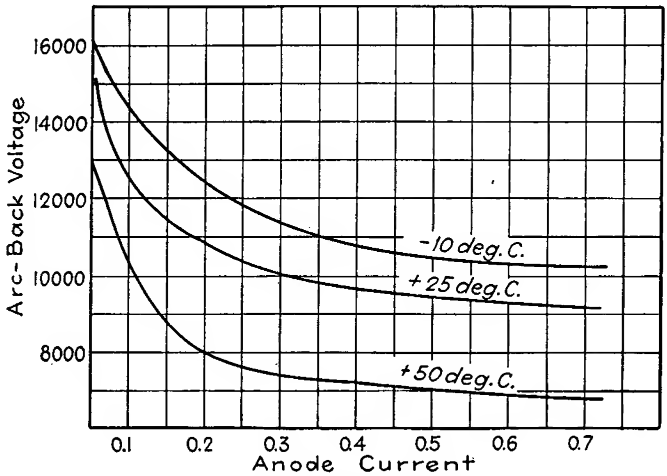


FIG. 79.—Arcback characteristics of a mercury-vapor diode.

20,000 volts. The tendency of the gas to "break down" under such high voltages depends on the pressure of the gas, which must be kept low enough to avoid ionization via the anode, but high enough to provide plentiful ionization via the cathode. In mercury tubes, in which the pressure varies with the temperature, the operating temperatures are restricted to a fairly narrow range, from about 20° to about 70°C., to insure the proper vapor pressure. The common name for reverse conduction (conduction of electrons from the anode to the cathode, the reverse of the normal direction) is *arcback;* it is one of the most annoying (and often hard to detect) aspects of gas-filled rectifier operation. A typical arcback voltage characteristic is shown in Fig. 79.

Gas-filled Diode Ratings.—The operating characteristics of gas-filled diode tubes are expressed in terms of several "ratings" which express the current-carrying ability and reverse-voltage limits. Since very large currents may be carried momentarily, it is customary to express two current-carrying values: the *maximum peak current* (which may be carried for short periods only) and the average current (averaged over a stated period) which can be carried continuously. Usually the maximum peak current is about four times as great as the average current. The *maximum peak inverse voltage* is the limiting negative voltage which may be placed on the anode during the nonconducting portion of the cycle. Its values range from 5000 to 20,000 volts, depending on the size and construction of the tube. The voltage drop between cathode and anode during conduction is also given, but since this value is not constant but varies slightly from tube to tube and with temperature and current flow, an average value only is given. Typical values are from 6 to 15 volts for mercury tubes.

Another highly important rating of gas-filled tubes is the *cathode-heating time.* If the cathode of the tube is not brought up to operating temperature before the anode voltage is applied, the required electron flow will not be available thermionically. The bombarding effect of the positive ions will nevertheless obtain electron regeneration (page 85) and the cathode will emit electrons by field emission and by secondary emission. This has a very harmful effect upon the thermionic properties of the cathode, lowering the emission level considerably if allowed to persist. To guard against these effects, it is necessary simply to have a plentiful supply of thermionically-emitted electrons available from the cathode before conduction begins. The time required to bring the cathode to the proper temperature varies with its construction and heating efficiency. Small tubes require about 5 sec., larger types take as long as 2 min., while the largest may require an hour. Automatic delay relays are often employed to withhold the anode potential for the required time. The current and voltage required to heat the cathode are also stated, and the permissible variations given. Usually a tolerance of ± 5 per cent is allowed.

High-pressure Gas-filled Diodes (Tungar-Rectigon Types).—A type of gas-filled diode used in battery-charging rectifier service

and other low-voltage applications is the Tungar or Rectigon tube. These tubes contain an inert gas (usually argon) at high pressure, *i.e.*, at pressures comparable with atmospheric pressure (760 mm.). The gas has two important effects; it ionizes and produces a current path of low-voltage drop, and it permits the operation of pure-metal filament (tungsten) at a very high temperature by inhibiting the tendency of the metal to evaporate. The filament provides large values of emission

TABLE V.—CHARACTERISTICS OF GAS-FILLED DIODE RECTIFERS (PHANOTRONS)

Type no.	Over-all length, in.	Over-all diameter, in.	Cathode type	Cathode, volts	Cathode, amp.	Maximum peak inverse, volts	Maximum peak, amp.	Maximum average, amp.
866, *PJ*-28, *WL*-866, 966, *C*-866	6⅝	2 7/16	Filament	2.5	5.0	7,500	1.0–2.0	0.25
866-*A*, *PJ*-28-*A*, *WL*-866-*A*, 966-*A*, *C*-866-*A*	6⅝	2 7/16	Filament	2.5	5.0	10,000	1.0–2.0	0.25
869-*A*, *PJ*-26-*A*, *WL*-869-*A*	14⅜	5 1/16	Filament	5.0	18.0	20,000	10.0	2.5
870, *FG*-42	25½	10¼	Heater	5.0	65.0	16,000	450.0	75.0
871, *WL*-871	4½	1 3/16	Filament	2.5	2.0	5,000	0.5	0.125
872, *FG*-19, *WL*-872, 972, *C*-872	8½	2 9/16	Filament	5.0	10.0	7,500	5.0	1.25
872-*A*, *FG*-19-*A*, *WL*-872-*A*, 972-*A*, *C*-872-*A*	8½	2 9/16	Filament	5.0	6.75	10,000	5.0	1.25
Tungar-rectigon Types								
289415	5⅝	2⅛	Filament	2.0	12.0	275	6	2.0
289414	6¾	2¼	Filament	2.2	18.0	300	19	6.0
766776	9½	3¾	Filament	2.5	27.0	225	47	15.0

because of its high temperature and large surface area. The high pressure of the gas makes reverse conduction possible at low voltage, so that the maximum inverse voltage is limited to about 400 volts. The arc starts when the applied voltage (making the anode positive) is about 12 volts, and the voltage drop during conduction is about 7 volts. The heater current in the filament is high, ranging from 12 to 30 amp., to obtain high operating temperature. The cathode-to-anode current ratings (average) range from 2 to 15 amp.

Table V gives the operating ratings of representative gas-filled diodes, of both the high- and low-voltage types.

33. Gas-filled Thermionic Triodes and Tetrodes (Thyratrons). The gas-filled diode suffers from the drawback we have already seen in the high-vacuum diode: Both are simply one-way current conductors, useful for rectifying alternating currents, but otherwise of no value in the control of power flow. The addition of a control electrode, or grid, to the gas-filled diode, increases its versatility almost as much as does the grid in the vacuum triode. But the action of the grid in the gas-filled triode, or thyratron, is very different from that in a vacuum tube.

The purpose of the grid in a thyratron is to control the *start* of the discharge. Once the discharge has begun, the grid is completely useless, so much so, in fact, that it might just as well be absent from the tube. The grid, in other words, is concerned with the initial breakdown of the gas, and its action must be interpreted in terms of the breakdown theory discussed in Chap. V (page 84). This theory states that gas or vapor breaks down and supports a self-sustained discharge when sufficient ions are produced to cause electron regeneration at the cathode. These ions are produced, in the main, by the impacts of electrons with neutral gas molecules, but only when the impacting electrons have sufficient energy to remove an electron from the molecule.

When the electrons leave the cathode emitter, they gather energy from their motion through the electric field near the cathode surface. If the strength of this field is great enough, the electrons gain the required ionizing energy between impacts, and ionization occurs. If, however, the field is weak, the electrons do not acquire the necessary energy, and no ionization occurs.

The grid structure in a gas-filled triode (Fig. 80), by its control of the electric field at the surface of the cathode, thus can prevent

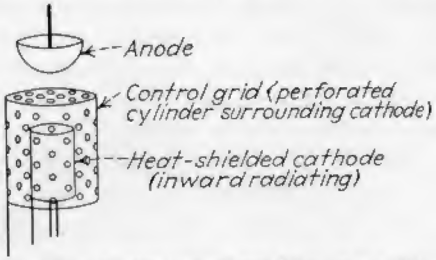


FIG. 80.—Element structure of gas-filled triode (thyratron).

the electrons from acquiring the necessary energy and prevent the start of the ionization. If the grid is maintained at a negative potential with respect to the cathode, it tends to reduce the electron energy. As it is made less negative (or more positive), the electron energy increases until the ionization level is reached.

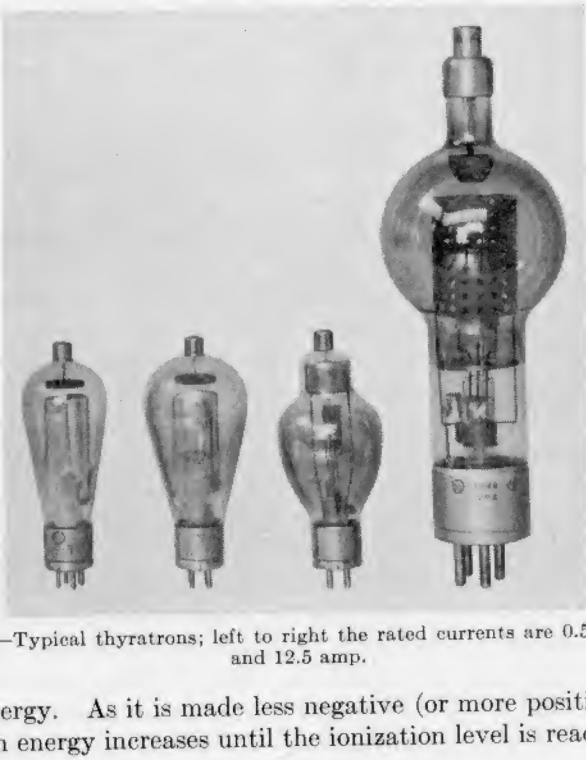


FIG. 81.—Typical thyratrons; left to right the rated currents are 0.5, 2.5, 2.5, and 12.5 amp.

When the discharge begins, the current which flows between cathode and anode is used to perform useful work in the anode circuit. But the current is quite different from that existing in a vacuum tube. In the first place, both positive ions and negative electron charges are present; the electrons, being of smaller mass, acquire the greater speed and hence carry the bulk of the current. The ions carry a current from anode to cathode, but this current is so small that its effect can usually be neglected so far as its effect on the external circuit is concerned. The presence of the ions is, on the other hand, of the greatest importance in the internal operation of the tube. The ionic effects are, in fact, the principal cause of the gas-filled tube's usefulness and at the same time of its limitations.

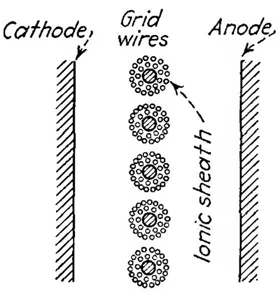


FIG. 82.—The ionic sheath on the grid wires, formed after conduction begins, prevents the grid from exercising control during the conduction period.

When ionization begins, the positive ions are attracted to the negative cathode where they form a concentration of positive charge which imparts the necessary electron energy to maintain the ionization. The action of the grid is, therefore, no longer necessary. This is fortunate, for at the same time the ions render the grid incapable of exerting any effect. If the grid is operated at a negative voltage with respect to the cathode, as it is in a majority of practical cases, some of the ions are attracted to the grid itself. Part of this ion flow to the grid causes a steady grid current; the remainder forms a sheath of ions near the grid which effectively shields the grid from the rest of the tube (Fig. 82). If the grid is made more negative, the thickness of the sheath increases and the effect of the grid on the electric field near the cathode (which might exercise some control on the cathode current) is thereby completely nullified. In some cases the grid is operated with a positive voltage, and in this case the ionic sheath may be very thin or lacking altogether, but the positive action of the grid is so small compared with the ionic concentration near the cathode that it has no effect on the cathode current. The grid is, then, useless as a control agent during the discharge. To stop the discharge it is necessary to remove the positive voltage from the anode, or at least to reduce it to a value

below the value (approximately 10 volts) necessary to maintain the discharge. The complete control of the current in a thyratron is thus dependent on both the grid and the anode, whereas in the vacuum triode it is dependent principally on the grid.

The important action of the gas-filled triode is the fact that when the grid voltage is brought to the proper value, a large current suddenly begins to flow between cathode and anode, and at the same time the voltage between cathode and anode is reduced sharply to the arc voltage drop. These sudden changes in current and voltage are the useful agencies which act on the load device which is connected in series with the anode. Inasmuch as these changes in current and voltage are under the control of an unrelated voltage applied to the grid, the tube is a relay device. But the control is of a very restricted nature. The current may be turned on by the action of the grid; but thereafter it cannot be turned off, nor changed in amount, by the action of the grid.

The Space-charge Conditions in Gas-filled Tubes.—It may be wondered why the current increases so suddenly when the gas discharge begins. The current flow comes, of course, from the cathode, since the extra electrons formed by ionization can participate in the cathode-to-anode current flow only so long as they are replaced by the cathode. Since the cathode is as good an emitter before the discharge starts as after, the change in current must be explained in terms of the electric field existing at the cathode surface.

Before the discharge starts, the only active electric charges available are the thermionically emitted electrons. These electrons form a space charge surrounding the emitter and limit the current flow in exactly the same way as in a vacuum tube. But when the discharge starts, positive ions are produced. These ions, being comparatively massive, move slower than the electrons, and the ionic charge thus has the opportunity of neutralizing the negative space charge of the electrons in the tube. The negative space charge near the cathode (and throughout the rest of the tube) is thereby reduced sharply, and the cathode current, freed from restraint, rises to whatever value is required by the external circuit. At the same time, the voltage between anode and cathode is reduced, by the reduction in space charge, to a very low value.

When the discharge has started, the potential distribution between cathode and anode assumes the shape shown in Fig. 83. This figure should be compared with Fig. 36 (page 65) which shows the potential distribution in a triode vacuum tube under similar circumstances. In the gas-filled tube, the major rise in voltage between cathode and anode occurs very close to the cathode, and the field strength at the cathode is correspondingly high. The voltage rise to the right (as shown in the diagram) is much more gradual, finally reaching a maximum near the anode. This "potential maximum," being the most positive part of the tube, acts as a storehouse for electrons, feeding them to the anode in proportion to the current requirement of the external circuit.

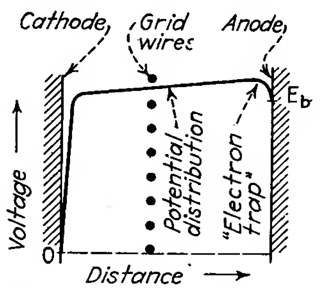


FIG. 83.—Potential distribution of a gas-filled triode.

Ionization and Deionization Time. In Chap. V, mention is made of the fact that a definite time is required to ionize the gas after the application of the necessary voltage, and that a similar interval is required to allow the gas to resume its neutral state after the arc-supporting voltage is removed. The time of starting, or *ionization time,* ranges from one to 50 microseconds, while the clean-up time, or *deionization time* may be as long as 1000 microseconds. Since the deionization time is longer, it is the principal limitation. An empirical expression for this time has been developed:

$$t = \frac{0.0012pI^{0.7}}{(E_g^{3/2} \times x)} \text{ sec.,} \tag{46}$$

where t is the time required, in seconds, for deionization in gas at pressure p baryes, when the arc current is I amp., the grid potential (with respect to the surrounding space) is E_g volts, and the grid-anode spacing is x cm.

The total time required for ionization and deionization is important because this time is not available for control purposes during the operation of the tube. If, for example, the controlling grid voltage is applied before deionization is complete, the grid will not exercise control. This fact limits the speed, or frequency, at which the tubes may be operated. At commercial

power frequencies (25 and 60 cycles per second), no trouble is usually experienced, but when more than 5000 tube operations per second are required, deionization difficulties are often serious. By proper design this limitation may be somewhat reduced. In a small mercury tube (the type 885) used in cathode-ray oscillography, full output is available at 15,000 operations per second, while some output may be obtained at 100,000 operations per second. The thyratron is thus a much more rapidly acting device than a mechanical relay, but it is considerably slower than the vacuum triode which will operate readily 30,000,000 times per second.

Grid Current in Thyratrons. Gas-filled Tetrodes.—Grid-current flow is much more troublesome in gas-filled tubes than in vacuum tubes. This results from the fact that both positive and negative charges are available, so that grid current will flow whether the grid is charged positively or negatively. At one particular value of grid voltage, equal numbers of positive ions and electrons will arrive at the grid in a given time, and at this voltage the grid current is zero. But this critical voltage value is of no practical value in reducing grid current since in operation the grid voltage must be changed from the critical value. Accordingly circuits in which thyratrons are used must be designed with grid current in mind. In particular this requires grid-control circuits capable of supplying considerably more power than is required in controlling vacuum triodes.

The grid current may be considered in two categories: that existing before the discharge begins and that existing during the discharge. Both types of current are important in circuit design. During the discharge, both ionic and electronic currents may flow to the grid, the predominant type being determined by the degree of positive or negative voltage applied to the grid.

The grid current existing just previous to the discharge is more important, since the grid is then exercising control, and the grid current produces a loss of power in the circuit supplying the control. The lower the grid current in this case, the easier the control from a small power source. The major part of the grid current during the control period results from emission from the grid. Small particles of oxide, which become dislodged from the cathode, may cling to the grid structure. The grid will then emit electrons when it is negative, and the resulting

grid current must be supplied by the control circuit. Grid emission may also start an arc between grid and anode, over which no control is available. Careful design may eliminate some of this trouble, by assuring low operating temperature and a small grid-surface area, properly protected from the cathode.

The shield-grid thyratron, a gas-filled tetrode, was designed principally to reduce grid current. In this type of tube a very much smaller control-grid structure is employed than in the triode thyratrons, and the grid current is correspondingly reduced. A typical metal-envelope tetrode structure is shown in Fig. 84. The extra electrode in the tetrode form permits changing the tube characteristics by changing the voltage applied to the screen grid. This allows compensation of changes which occur as the tube ages or when an old tube is replaced by a new one. This is an important point, since it is difficult to manufacture gas-filled tubes within the tolerances possible in vacuum tubes.

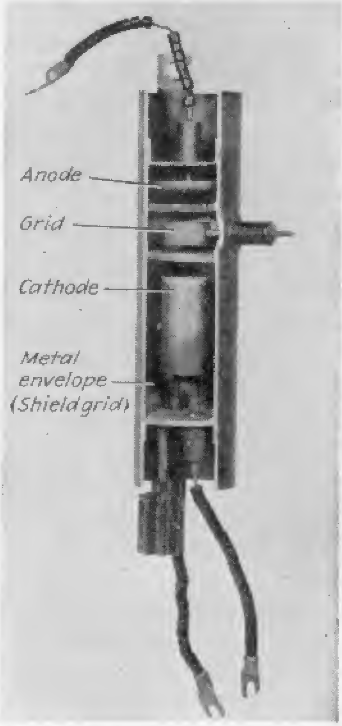


FIG. 84.—Cut-away view of a typical shield-grid thyratron.

Operating Characteristics of Thyratrons.—Since thyratrons are gas-filled diodes in which one or two grids have been inserted for control purposes, all the operating characteristics stated for gas-filled diodes (page 138) apply equally well to the thyratron family. The maximum-peak anode current, the average anode current (and the time of averaging) the maximum-peak inverse voltage, the cathode-heating time, and the cathode-heating current and voltage must all be specified for triodes and tetrodes. One other voltage, which applies to the anode-cathode circuit, must also be specified. It is the maximum-peak *forward* voltage, the value of positive anode voltage against which the grid can withhold the start of the arc. Forward peak voltages of from 500

to 10,000 volts are applicable to many thyratrons, depending on the effectiveness of the grid in shielding the cathode from the effect of the anode. The forward voltage, after conduction begins, is, of course, the usual arc drop of from 6 to 30 volts, depending on the construction of the tube and the type of gas employed in it. In mercury tubes these ratings, especially the arc-voltage drop, vary with temperature, so that it is necessary

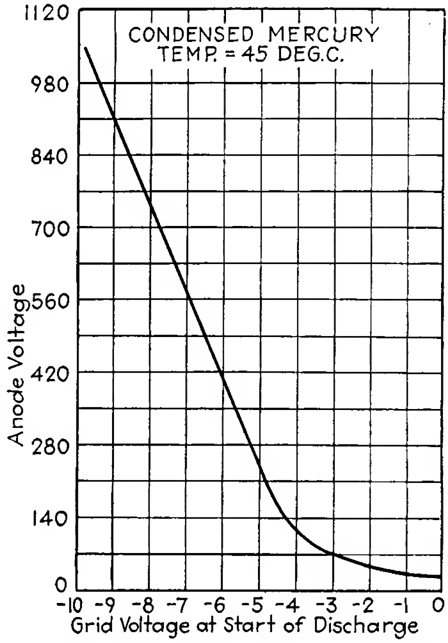


Fig. 85.—Grid-control starting characteristic of a typical negative-control thyratron (Type KU-676).

to specify the temperature of the condensed mercury within the tube, which applies with the given ratings. The neon, argon, or helium tubes usually are fairly stable with respect to temperature, since the mean free path between molecules remains essentially constant.

Grid-control Characteristics.—The action of the grid in controlling the discharge is expressed by the voltage (between grid and cathode) at which the discharge will start. Since the anode voltage also contributes to the field at the surface of the cathode at the start of the discharge, both grid and anode voltages must

be specified. In general, the higher the positive anode voltage, the more readily will the discharge start and the more negative must the grid be to restrain the current flow.

This relationship between grid and anode voltages required to start firing is expressed graphically in a control characteristic, in which voltage between grid and cathode is plotted horizontally against the anode voltage vertically. Any point on the

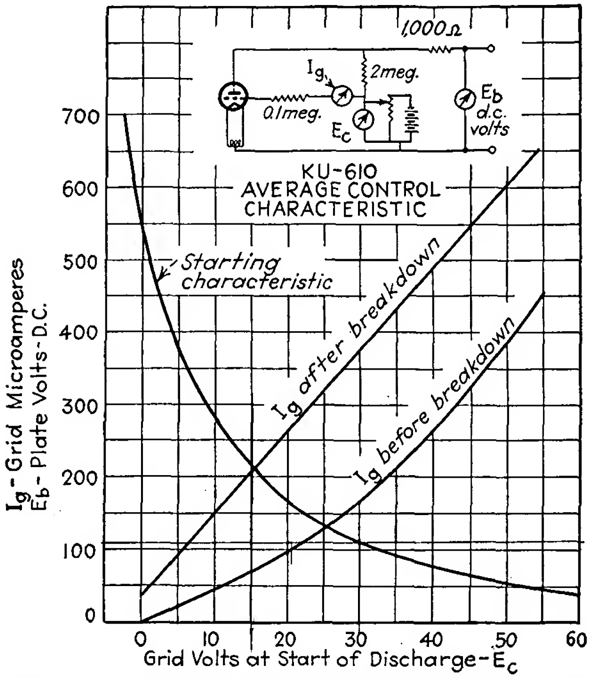


FIG. 86.—Starting characteristic of a positive-control thyratron. Grid current values before and after conduction begins were taken in the circuit shown.

control curve specifies a combination of grid and anode voltage at which "firing" (*i.e.*, the start of the discharge) will take place. If the anode voltage is lower than this value, conduction will not begin unless the grid voltage is raised, or if the grid voltage is more negative, conduction will not begin until he anode voltage is raised. The grid-control characteristic is very useful in plotting the control locus (page 260) when alternating voltage is applied to both grid and anode, that is, when various combinations of grid and anode voltages are presented in quick succession.

Control characteristics are divided into two types, positive and negative. The first class contains tubes designed for

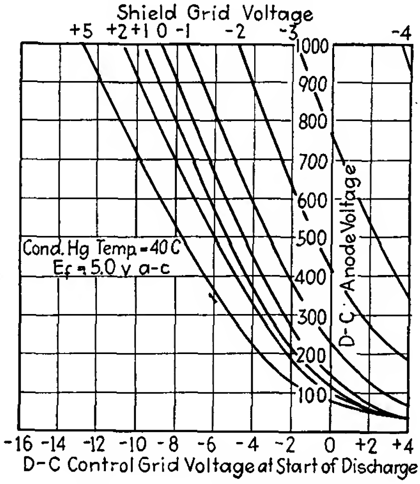


FIG. 87.—Starting characteristic of a shield-grid thyratron (Type FG-105).

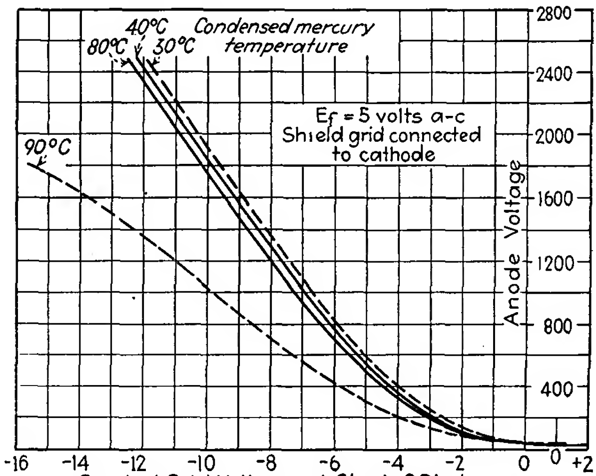


FIG. 88.—Variation of control characteristic with temperature of condensed mercury (Type FG-172).

operation with their grid *positive* over all or part of the control range. Such tubes draw large grid currents from the control circuit. The negative types (usually mercury-vapor-filled) are restricted to operation with the grid negative with respect to

cathode. The latter class of tube, because of the mercury-vapor content, usually has a considerable temperature coefficient. In fact, it is customary to plot three or four control characteristics, for various temperatures within the allowable temperature range. Typical examples of these control characteristics are given in Figs. 85 to 88.

The grid-current characteristic is also specified, especially for the positive-control tubes. Often a complete control char-

TABLE VI.—CHARACTERISTICS OF THYRATRONS
Mercury-vapor and Inert-gas-filled Types

Type no.	Over-all length, in.	Over-all diameter, in.	Cathode type	Cathode volts, amp.	Maximum peak forward, volts	Maximum peak inverse, volts	Maximum peak, amp.	Maximum average, amp.	Grid type	Deionization time, microseconds
885	4 1/4	1 9/16	Heater	2.5,1.4	350	350	0.300	0.075	Neg.	1,000
FG-17	6 5/8	2 7/16	Filament	2.5,5.0	2,500	2,500	2.0	0.5	Neg.	1,000
FG-29	14 5/8	5 1/16	Heater	5,17.5	3,500	3,500	75	12.5	Neg.	1,000
FG-33	7 1/2	3.	Heater	5,4.5	1,000	1,000	15	2.5	Pos.	1,000
FG-37	7 1/2	3	Heater	115,0.2	1,000	1,000	15	2.5	Pos.	100
FG-41	17 11/16	5 1/16	Heater	5,20	10,000	10,000	75	12.5	Neg.	100
FG-118	17 5/8	5 1/16	Heater	5,20	10,000	10,000	75	12.5	Pos.	
FG-57	7 1/4	3.	Heater	5,4.5	1,000	1,000	15	2.5	Neg.	1,000
FG-65	4 5/8	1 3/16	Filament	2.5,2.0	1,000	1,000	0.5	0.125	Neg.	1,000
FG-67	7.	3.	Heater	5.0,4.5	1,000	1,000	15	2.5	Neg.	100
FG-81*	6 5/8	2 7/16	Filament	2.5,5.0	180	180	2.0	0.5	Neg.	1,000
FG-178*	4 5/8	1 3/16	Filament	2.5,2.25	310	310	0.500	0.125	Neg.	1,000
FG-27-*A*	7 1/4	3.	Filament	5.0,4.5	1,000	1,000	10	2.5	Neg.	1,000
KU-623*H*	14 3/8	5 1/8	Heater	5.0,20.0	5,000	10,000	10	2.5	Neg.	
KU-627	7	2 7/16	....	2.5,6.0	1,250	2,500	2.5	0.64	Neg.	
KU-828	9 1/4	3 1/4	....	5.0,11.5	1,250	2,500	8	2	Neg.	
KU-633	5	1 9/16	....	2.5,2.0	500	1,000	0.5	0.125	Neg.	
KU-634	9	3 13/16	....	5.0,11.5	7,500	7,500	5.0	1.25	Neg.	
KU-676	12 1/4	3 7/8	....	5.0,11.5	1,000	1,000	40	6.4	Neg.	
KU-610*	6 1/2	2 7/16	....	2.5,6.5	750	1,500	0.40	0.10	Pos.	
KU-635*	9	3 1/4	....	5.0,14.0	350	350	4.0	1.0	Neg.	
KU-636*	7	2 7/16	....	2.5,7.0	350	350	0.4	0.1	Neg.	
Double-grid (shield-grid) Types										
FG-95	7 15/16	4 1/2	Heater	5.0,4.5	1,000	1,000	15	2.5	Neg.	1,000
FG-98*	6 3/4	4	Filament	2.5,5.0	180	180	2.0	0.5	Neg.	1,000
FG-97	6 3/4	4	Filament	2.5,5.0	1,000	1,000	2.0	0.5	Neg.	
FG-105	11 1/4	3 13/16	Heater	5.0,10.0	1,000	1,000	40	6.4	Neg.	1,000
FG-154*	7 15/16	3	Filament	5.0,7.0	500	500	10	2.5	Neg.	1,000
FG-172*,†	10 3/4	3 5/8	Heater	5.0,10.0	1,000	1,000	40	6.4	Neg.	1,000

* Inert-gas-filled types. All others mercury-filled.

† Metal envelope.

acteristic is given of the tube in a typical circuit, including the effects of the recommended grid-current-limiting resistor, as in Fig. 86.

The range of gas-filled triodes, in size and power-handling capacity, is fully as great as that of the gas-filled diode rectifiers. The inverse-voltage ratings are as high as 10,000 volts with forward-voltage ratings of approximately the same magnitude. Table VI gives a listing of commonly used gas-filled triode and tetrode tubes with their ratings.

34. Pool-type Diodes and Triodes. The Ignitron Principle.— Gas-filled electron tubes employing mercury-pool cathodes are at once the oldest and the newest of the gas-filled family. The Cooper-Hewitt lamp, an early gas-discharge device, and the Y-type rectifiers used for years in telephone central-offices are familiar diode forms of this type. The newer ignitron, a triode form, is in many ways the most promising of all the electron tubes available for heavy-duty service.

The use of mercury as a cathode is based on several advantages which it enjoys over the oxide surface used in the thermionic gas-filled tubes. The oxide-coated cathode requires time to arrive at operating temperature, hence the cathode-heating power must be applied continuously, a wasteful procedure when the tube is operating in intermittent service. The mercury-pool cathode requires no heating power at all, and is ready to emit at any time. In the second place, the overload capacity of the oxide emitter is limited. If excessive currents are drawn from the oxide, in excess of the safe thermionic limit, the high field at the surface will cause disintegration of the surface under ionic bombardment. The mercury-pool cathode, on the other hand, has an almost unlimited overload capacity; as many cathode spots form as are required to supply the required emission. The mercury pool is thus both convenient and flexible.

The presence of the mercury pool in the cathode implies that the rest of the tube will contain mercury vapor at the vapor pressure corresponding to the temperature of the pool. The active current-carrying agent in the tube is thus necessarily mercury vapor. This is not a disadvantage since mercury vapor is a very desirable carrier. Being a heavy material, its ions are sluggish and thus very effective reducers of the electronic space charge. The ionization and resonance potentials of mercury

vapor (page 83) are lower than those of many other gases, resulting in a low arc voltage drop, important from the standpoint of efficiency and resulting in less energetic bombardment of the cathode. Finally the presence of the liquid mercury in the tube provides an inexhaustible supply of vapor, whereas gases are apt to disappear, or "clean up," during the life of the tube by combining chemically or physically with the electrodes.

The high current output of the mercury-pool cathode and its high overload capacity fit it for heavy-duty service. Pool-type tubes are designed therefore for anode currents of from 10 amp. up to 5000 amp. For currents less than 5 amp. the operation of the electron-emitting cathode spot is apt to be somewhat erratic.

One disadvantage of the mercury pool is the fact that the cathode spot is not formed spontaneously. It is necessary to apply voltage directly across the liquid metal itself and to draw an arc by removing one of the electrodes from the pool. The arc so formed starts the cathode spot, which will thereafter persist until the anode voltage is removed. If the anode voltage is removed only momentarily, the spot may "carry over" to the next cycle, but reignition of the spot at the beginning of every cycle may be necessary. At commercial power frequencies, if both halves of each alternating cycle are active—as they are in the full-wave type of rectifier—the spot ordinarily persists from cycle to cycle. The necessity of initiating the spot at the beginning of operation still remains, however, and many ingenious devices have been employed to perform this duty. The Y-type rectifier is started by tipping the entire tube until the mercury pool covers an auxiliary starting electrode. The Hewitt lamp is started by the application of a very high voltage to an external electrode, which induces the starting current in the pool. In many tubes an auxiliary "keep-alive" electrode is supplied continuously with power to maintain the spot regardless of the anode voltage.

The most elaborate form is the tank type of rectifier (Fig. 89), containing a single mercury pool at the base which acts as the common cathode for several anodes above which act, successively, to draw the load current from the cathode. Each anode is connected to a separate phase of the load circuit, hence the sequential action of the anodes keep the load current at a

high value at all times, so that the cathode spot never has an opportunity to fail. An auxiliary starting electrode is available for beginning operation and automatically reigniting the arc after an interruption of service. The capacity of tank rectifiers runs from 100 to 5000 amp. per anode. The large size of the tanks and the consequent possibility of leaks make desirable an

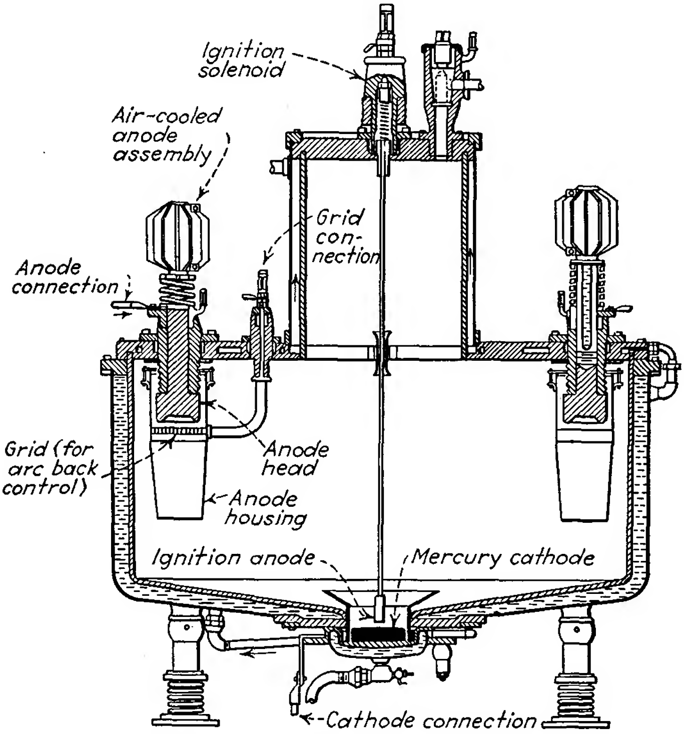


FIG. 89.—Multi-anode tank rectifier. (*Cf.* Fig. 1, page 5.)

auxiliary pumping system to maintain the vacuum within the envelope, in the form of an automatic pump which is turned on when the pressure exceeds a few thousandths of a millimeter. Tank rectifiers are now usually provided with grid structures near each anode which permit external control of the timing of each conduction cycle; these tanks are quite similar in action to the ordinary thyratron.

Triode Pool-type Tubes. The Ignitron.—The great advantage of the high current and overload capacity of the mercury-pool

has been combined with the control versatility of the thyratron in the *ignitron.* The ignitron (Figs. 90 and 91) contains a control electrode, or ignitor, which is partially immersed, permanently, in the pool cathode. By applying voltage between this starting electrode and the pool, the cathode spot is formed and the tube caused to start operation at any desired time, so long as sufficient positive voltage is applied to the anode. The ignitron differs from the thyratron in that it will not start until forced to do so, whereas the thyratron remains nonconducting only so long as it is prevented from starting. This is an important distinction in many classes of service.

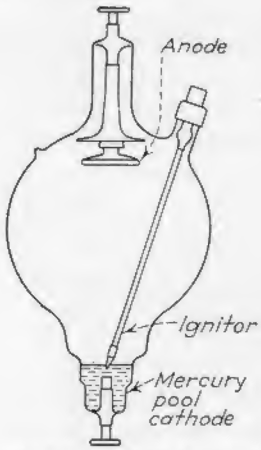


FIG. 90.—Arrangement of ignitron element structure.

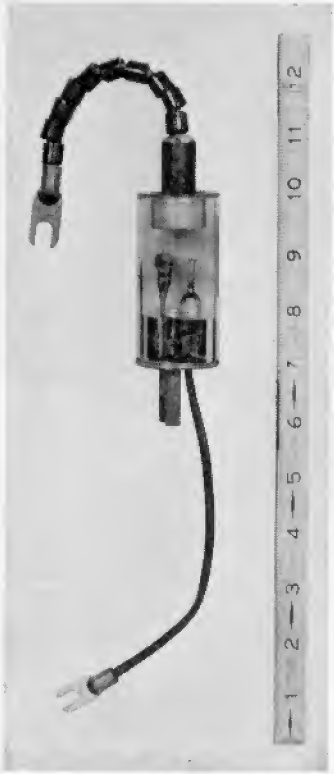


FIG. 91.—This 3-in. ignitron will conduct several hundred amperes. The point-shaped element (left center) is the ignitor.

The ignitor electrode is usually a rod of high-resistance refractory material, commonly boron carbide, shaped in the form of a conelike solid of revolution. When this rod is immersed in the mercury, the surface tension of the mercury produces a very small separation between rod and mercury below the surface of the pool. Part of the voltage applied between the ignitor

and the mercury pool is thus caused to act across a very thin vapor-filled space, and the electric field in that space is correspondingly high. The field ionizes the vapor and the arc starts. The cathode spot, very minute at first, immediately increases in size and rises to the surface of the pool. Thereafter the anode attracts the required electron emission from the spot and the anode current flows. The ignitor then has no control, of course, until the arc has been extinguished by the reduction of the anode voltage to the extinction point. The time required

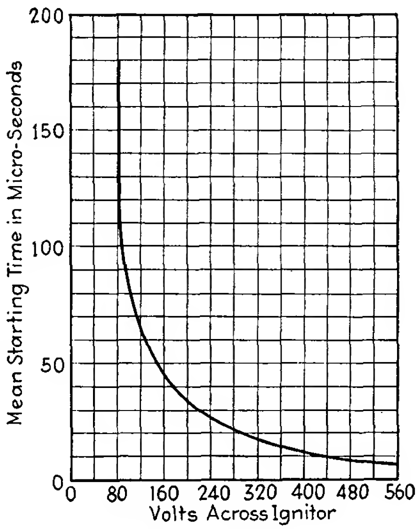


FIG. 92.—At high ignition voltages, the ignitron starts in less than 10 microseconds. Above is the starting time curve of the FG-253.

for the ignition of the arc by the ignitor electrode is very small usually less than 50 microseconds.

The ignitor principle, starting the arc only when needed, has another advantage over the keep-alive diode form of tube in which the arc is kept alive continuously. The keep-alive current used for maintaining the cathode spot in the diode tube tends to produce ionization during the inverse half of each cycle, from which arcbacks can develop readily. The ignitron tube, size for size, is thus much less prone to arcback than is a diode pool-type tube with a keep-alive auxiliary.

The current and voltage which must be applied to the ignitor to start the arc are rather large. The starting voltage

is from 100 to 250 volts and the starting current from 3 to 40 amp. But since the starting power need be applied only for about 50 microseconds, the average power (in 60-cycle service)

TABLE VII.—CHARACTERISTICS OF IGNITRON TUBES

Type no.	Over-all length, in.	Over-all diameter, in.	Maximum ignition, volts (instantaneous)	Maximum ignition, amp. (instantaneous)	Maximum peak, inverse volts†	Maximum peak, amp.†	Maximum average, amp.†	Maximum forward volts†
KU-637	11⅝	5⅛	150	15	750	1,000	20	750
KU-639	20⅝	10⅛	100	24	750	1,250	50	1,750
KU-671	23½	10⅛	100	25	750	2,000	64	750
KU-651*	15.	4¼	100	25	360	4,000	85	360
FG-139	15¼	4¼	150	15	900	1,125	12.5	900
FG-179	24	10¼	...	..	1,000	5,000	25	1,000
FG-179-*A*	23¾	6¼	...	..	900	2,500	35	900
FG-194*	5⅜	1⅝	250	10	900	800	2.25	900
FG-235-*A*	15	4¼	...	..	900	4,000	75	900
FG-238-*A*	....	....	250	40	2,100	1,800	200	2,100
FG-253*	9⅛	2¼	150	15	900	600	4.0	900
FG-258-*A*	18.	6	250	40	900	10,000	250	900
FG-271	12.	2¾	250	15	900	2,100	40	900

* Metal envelope.

† Ratings given under each type are maximum allowable and are not corresponding ratings.

required for ignition may be only a fraction of a watt. Since the control current is high, it is usual to employ a thyratron tube as the source of control current in applications where precision of timing is required. A simple rectifier tube, operating on the inverse voltage of the ignitron itself, is also commonly used to supply the ignition current. A condenser discharge current may also be used.

The ratings of ignitron tubes (see Table VII) are similar in form to those of the thyratrons. The peak current is usually much higher than the average current, often 50 times as great. Typical values are 20 amp. average, 1000 amp. peak. The inverse and forward (nonconducting) voltages are fairly low, usually not higher than 1,000 volts. The arc voltage drop (during conduction) is, as usual, about 15 volts. The current capacity is limited not by the cathode, which will handle any current required, but by the heat generated in the arc. The type of cooling employed thus has a very great effect on the current

ratings. A typical ignitron has a current capacity four times as high under forced cooling as under ordinary convection cooling.

Ignitron tubes are widely used in welding-control service because they combine heavy current capacity with a high degree of precision of timing. Typical circuits of this type are described in Chap. XIII.

35. Cold-cathode Gas-filled Triode Tubes.—At the opposite extreme from the mercury-pool cathode in emitting capability is the cold cathode, a simple metallic electrode which may or may not be coated with a barium compound to lower its work function. No heating current is employed, and consequently there is no steady power drain for this purpose, an advantage in intermittent service and in locations where power is at a premium. The emission of electrons from the cold cathode is a very complex phenomenon, being due to field emission, secondary (bombardment) emission, with traces of thermionic and photoelectric emission. A comparatively high field strength is thus required to produce the necessary emission, and the emission current is in any event small. A typical cold-cathode thyratron tube (the Westinghouse KU-618, a four-element neon-filled tube) has a cathode-to-anode voltage drop (during conduction) of 180 volts, supplies anode currents of 0.015 amp. average, 0.100 maximum peak, and will withstand a peak inverse voltage of 800 volts.

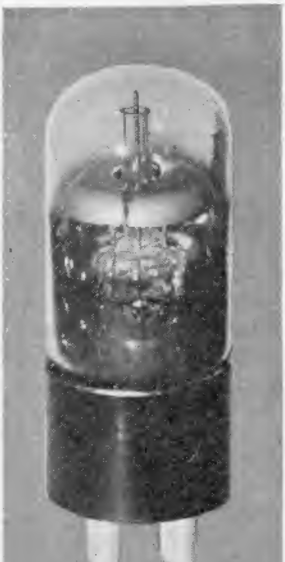


FIG. 93.—Cold-cathode neon-filled relay tube (Type 313-*A*).

Another form of cold-cathode relay tube of considerably different construction is the Western Electric 313-*A*, shown in Fig. 93. As the photograph shows, there are two semicircular elements either of which may act as the cathode, the other as the control element, and a wire partially shielded in glass which acts as the anode. The small separation between cathode and control elements reduces the voltage necessary to start the discharge to

about 70 volts. An active barium coating on these electrodes contributes to this low voltage and makes the cathode-anode voltage drop small, about 75 volts. The maximum inverse voltage between anode and cathode is about 175 volts. The tube is used for bell-ringing control in subscribers' telephone equipment.

Several other forms of cold-cathode tubes are used for control purposes, but they follow the general principles employed in the KU-618 and the 313-A. All are characterized by high values of control and arc voltage, with comparatively low values of inverse voltage. They are used principally in low-current applications, where the inconvenience or power requirements for a cathode-heating supply cannot be tolerated.

Problems

1. A typical gas-filled diode mercury-rectifier has an arc-voltage drop of 12 volts (assumed constant throughout conducting cycle) at 35°C., and starts conduction when the applied voltage is 16 volts. Show by graphical construction what average current is carried through this tube when the load resistance is 1000 ohms and the applied voltage (across tube and load) is 110 volts r-m-s, pure sine-wave alternating current. (HINT: Employ method used in Prob. 3, at the end of Chap. VI, page 132.)

2. Using the thyratron-control characteristic at 80°C. shown in Fig. 88, prepare a plot of the grid voltage (plotted against time) necessary to prevent firing of the tube throughout a full cycle of 800-volt r-m-s alternating voltage applied between the anode and cathode of the tube. If the applied grid voltage is −5.0 volts, show at what point the conduction will begin and estimate the average current passed through a 1000-ohm load if the conduction thereafter continues to the end of the cycle. Estimate as the average current if the applied voltage on the grid is −7 volts. Neglect arc-voltage drop.

3. Repeat Prob. 2 for an operating temperature of 90°C.

4. From the grid-current characteristic shown in Fig. 86, determine the maximum grid-current flow to the thyratron grid, between discharges, of a 5.0-volt r-m-s alternating control voltage applied to the grid (+15 volts d-c bias). The applied anode voltage (with respect to cathode) is 110 volts, r-m-s, in phase with the grid voltage.

Bibliography

KOLLER, L. R.: "Physics of Electron Tubes," McGraw-Hill Book Company, Inc., New York, 1937, Chaps. X, XI.

MACARTHUR, E. D.: "Electronics and Electron Tubes," John Wiley & Sons, Inc., New York, 1936, Chap. VII.

DOW, W. G.: "Fundamentals of Engineering Electronics," John Wiley & Sons, Inc., New York, 1937, Chaps. XIX, XXI.

MARTI and WINOGRAD: "Mercury Arc Power Rectifiers," McGraw-Hill Book Company, Inc., New York, 1930.

PRINCE and VOGDES: "Principles of Mercury Arc Rectifiers and Their Circuits," McGraw-Hill Book Company, Inc., New York, 1927.

HULL, A. W.: Gas-filled Thermionic Tubes, *Trans. A.I.E.E.*, **47**, 753 (1928).

HULL, A. W.: Hot Cathode Thyratrons, *Gen. Elec. Rev.*, **32**, 213 and 390 (1929).

HULL, A. W.: Characteristics and Functions of Thyratrons, *Physics*, **4**, 66 (1933).

HULL and LANGMUIR: Control of an Arc Discharge by Means of a Grid, *Proc. Nat. Acad. Sci.*, **51**, 218 (1929).

LIVINGSTON and MASER: Shield-grid Thyratrons, *Electronics*, **7**, 114 (1934).

SLEPIAN and LUDWIG: A New Method of Starting an Arc, *Elec. Eng.*, **52**, 605 (1933).

KNOWLES, D. D.: The Ignitron, A New Controlled Rectifier, *Electronics*, **6**, 164 (1933).

CHAPTER VIII

PHOTOSENSITIVE TUBES AND CELLS

Introduction.—There are three forms of photosensitive device which are commonly called "electronic," the photoemissive, the photovoltaic, and the photoconductive. *Photoemissive tubes* make use of photoelectric emission (page 43), the current flow through them being provided by the emission of electrons from a photosensitive cathode surface. Since the current flow in these tubes takes place in a gas or in a vacuum, they are true electronic tubes, in accordance with our definition of the word. The other two types are not electronic in that the current flow in them takes place in a solid or liquid. The *photovoltaic cells* convert the light energy falling on them directly into an electric current. The *photoconductive cells* contain a semiconducting solid substance whose resistance to electric-current flow decreases as the amount of light falling on it increases. All three types of device depend basically on the fact that light falling on any substance tends to free electrons from that substance.

36. Photoemissive Electron Tubes. Classifications and Characteristics.—Photoemissive electron tubes, commonly called phototubes, contain a cathode whose surface is treated to produce a low work function and hence emits electrons when light falls on the surface. A second electrode acts as an anode to collect the emitted electrons. The cathode and anode are sealed in a glass envelope which is either highly evacuated or filled with inert gas at very low pressure. The gas-filled tubes are more sensitive than the vacuum type but are not so constant in their characteristics. The tubes are classed according to the type of cathode employed, and according to the gas content, as gas-filled or high-vacuum tubes. Various differences in the structure of the cathode and anode, and in the type of glass employed in the envelope are available for specialized purposes.

Characteristics of Photoemissive Cathodes.—The cathode of a phototube determines two of its most important characteristics,

TABLE VIII.—CHARACTERISTICS OF PHOTOEMISSIVE TUBES

Vacuum Types

Type no.	Cathode material	Average sensitivity, μa. per lumen	Maximum anode voltage, volts	Maximum anode current, μa.	Cathode window area, sq. in.	Spectral limits, Å.	Wave lengths of maximum sensitivity, Å.
PJ-22	Caesium-oxide-silver (visible, infrared)	14	200	20	0.9	3,000 to above 8,000	3,500 7,500
FJ-76	Sodium (ultraviolet range)	12	200	50	2.0	2,000 to 5,000	2,900
FJ-114	Caesium-oxide-silver (visible, infrared)	35	200	20	0.6	3,000 to 10,000	3,500 8,800
917 } 919 }	Caesium-oxide-silver (visible, infrared)	20	500	30	0.9	below 5,000 to 12,000	7,900
922	Caesium-oxide-silver (visible, infrared)	20	250	30	0.4	below 4,000 to 11,000	8,000
WL-770	Caesium-magnesium (near ultraviolet)	0.75	500	6	2.4	2,750 to 7,000	3,300
SR-50 } *WL*-734 }	Caesium-oxide-silver (visible, infrared)	15	500	22	1.1	2,900 to above 8,000	3,500 7,000
SR-51	Caesium-oxide (violet)	15	500	22	1.1	2,800 to 9,000	4,000
SR-53	Caesium-oxide (visible, infrared)	25	500	22	1.1	2,900 to above 8,000	3,500 7,000
WL-774	Tungsten on nickel (ultraviolet)	0.001	500	..	.75	1,700 to 2,700	2,400
WL-773	Thorium-on-nickel (ultraviolet)	0.10	500	..	1.33	2,500 to 3,650	3,000
WL-767	Titanium-on-nickel (ultraviolet)	0.02	500	..	1.33	2,500 to 3,200	2,925

Gas-filled Types

Type no.	Cathode material	Average sensitivity, μa. per lumen	Maximum anode voltage, volts	Maximum anode current, μa.	Cathode window area, sq. in.	Spectral limits, Å.	Wave lengths of maximum sensitivity, Å.
PJ-23	Caesium-oxide-silver (infrared, visible)	50	90	20	0.9	3,000 to above 8,000	3,500 7,500
868	Caesium-oxide-silver (visible, infrared)	55	90	20	0.9	3,000 to 11,000	3,500
918	Caesium-oxide-silver (visible, infrared)	110	90	20	0.9	below 4,000 to 12,000	8,000
920	Caesium-oxide-silver (visible, infrared)	75 (per unit)	90	15	0.31 (per unit)	below 4,000 to 11,000	7,500
921	Caesium-oxide-silver (visible, infrared)	100	90	20	0.4	below 4,000 to 12,000	7,000
WL-735 } *SK*-60 }	Caesium-oxide-silver (visible, infrared)	60	90	22	1.1	2,900 to above 8,000	3,500 7,000
SK-63	Caesium-oxide-silver (visible, infrared)	125	90	22	1.1	2,900 to above 8,000	3,500 7,000

the *luminous sensitivity* and the *spectral response.* The luminous sensitivity is the measure of the amount of current available from the cathode when a given amount of light falls upon it; it is expressed in microamperes per lumen. The values range from a fraction of a microampere (μa.) to over 100 μa. per lumen of light, depending on the composition of the cathode surface and the gas content of the tube. Table VIII shows the sensitivities of several commonly used cathode surfaces. The most sensitive is the caesium-oxide-on-silver combination, described on page 44. As is to be expected, the sensitivity of the surface increases as the work-function energy decreases.

In general the currents available from photoelectric cathodes are very small, usually not more than a few microamperes. The available emission current depends directly, of course, on the number of lumens of incident light. If all the light from a 32-candlepower automobile head lamp ($32 \times 12.6 = 403$ lumens) were to be concentrated on a caesium-oxide cathode (20 μa. per lumen), the total current would be 8000 μa., or 8 ma., but such an intense current would damage the photosensitive surface. The safe limit for such a cathode is not more than 40 μa. corresponding to 2 lumens of light.

The small values of current available from photoemissive cathodes impose a severe limitation on the types of service in which phototubes can be used. The anode current is so small that it cannot be made to actuate any device directly except a very sensitive meter, hence phototubes are universally used in conjunction with amplifier tubes (vacuum or gas-filled triodes) capable of converting the feeble anode current into a current capable of controlling a relay or some similarly rugged device.

The value of the sensitivity of the cathode depends, of course, on the color of the light which falls on it. The industry has standardized on a tungsten-lamp source operated at 2870°K. as an arbitrary basis for comparing the sensitivities of different tubes.

The *spectral response* of a cathode expresses the relative amount of photoelectric current produced by light of different colors. It is represented by means of a plot, the current in microamperes per microwatt of radiant energy against the wave length or frequency of the light, as shown in Fig. 94. It is necessary that the strength of the light be the same for every

color, so the current is expressed as microamperes per lumen or per microwatt of incident light (one lumen = 0.0016 watt for green light). The spectral-response curves are, in general, partial

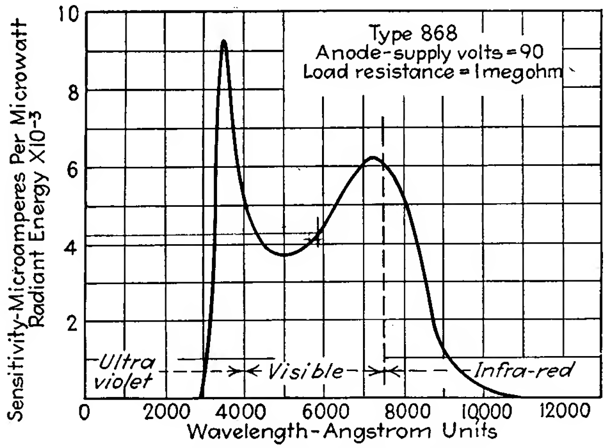


FIG. 94.—Spectral-response curve of a caesium-oxide-silver cathode. Note the response in the infrared region.

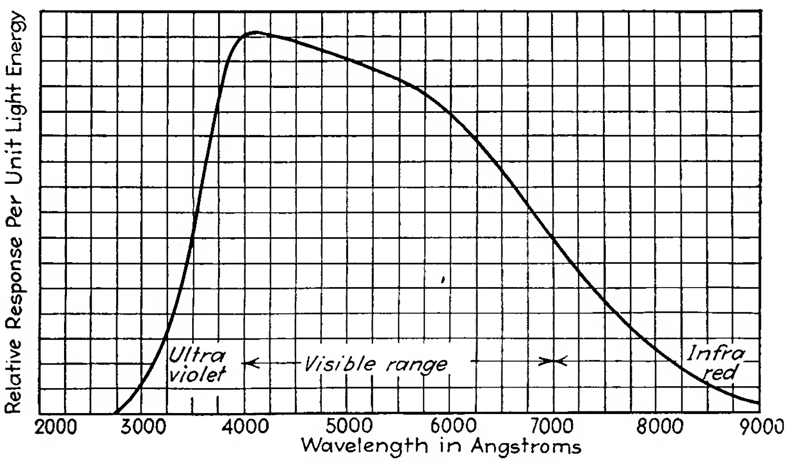


FIG. 95.—Violet-sensitive cathode.

to one or two regions of the spectrum, as shown in Figs. 94, 95, and 96; no cathode substance has as yet been found which responds equally well to all colors.

In general the cathodes with the lowest work functions have the best response in the red and infrared regions, (in which a large

portion of the light from a tungsten lamp resides) while those of high work function (tungsten, for example) respond well in the ultraviolet region. The glass of the envelop is also of importance. Most glass transmits visible light well, but infrared and ultraviolet may be greatly weakened by passing through even a thin layer of ordinary glass. For research with the ultraviolet, phototubes are sometimes made with fused-quartz envelopes, since this substance transmits the high-frequency light well. Likewise, very thin windows (see Fig. 101*E*) are

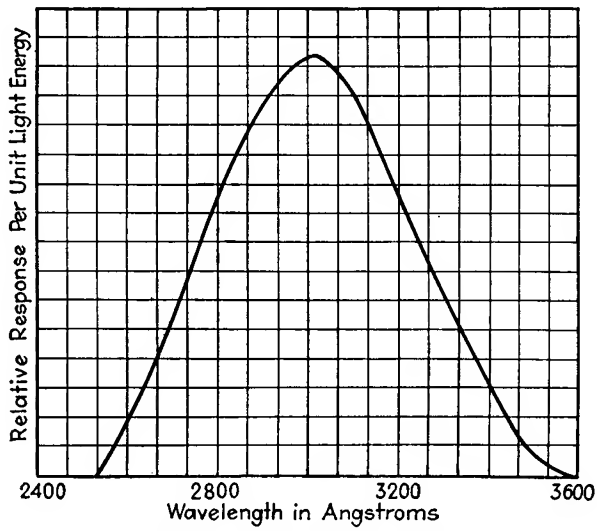


FIG. 96.—Ultraviolet sensitive cathode (thorium on nickel).

sometimes employed to increase the amount of light transmitted to the cathode.

Cathode-to-anode Characteristics.—The current which flows through a phototube is dependent not only on the emitting characteristics of its cathode but also upon the position of the anode with respect to the cathode. The anode of most phototubes is a very simple affair, either a straight wire or a ringlike structure supported so that it does not obstruct the cathode any more than is necessary. The current-voltage relationship of a vacuum-type tube follows the $\frac{3}{2}$-power law, in accordance with Langmuir's theorem which states that the current between electrodes in a vacuum increases as the $\frac{3}{2}$ power of the applied voltage regardless of the shape and separation of the electrodes. When

the tube is gas-filled, no such simple relationship exists, due to the effect of the ionization on space charge, but it is still possible, of course, to express the relationship in this case by a curve of current versus voltage plotted from actual measurements.

In Figs. 97 and 98 are given typical current-voltage curves of vacuum and gas-filled tubes. It will be noticed in the vacuum case that the $\frac{3}{2}$-power law holds for very low values of voltage (below 5 volts), above which saturation sets in, the level of saturation current depending on the amount of light falling on

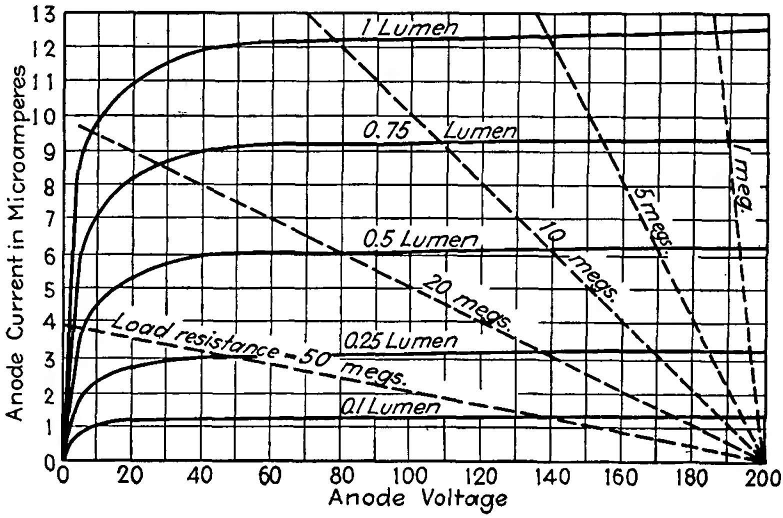


FIG. 97.—Anode-current versus anode-voltage characteristic of a vacuum-type phototube (Type PJ-22).

the phototube cathode, since the saturation current represents the total emission from the cathode. In the gas-filled case, however, the saturation condition persists over a short range of voltage only. As the voltage increases, the tendency of the gas to ionize increases also and the current therefore increases to very great values. If too high a voltage is applied, the ionization will result in electron regeneration at the cathode, and a self-sustained discharge will occur. The bombardment incident to this discharge, if allowed to persist, will greatly increase the work function of the surface, thus lowering its photosensitivity. Precautions must be taken, therefore, to prevent excessive voltage from being applied to a gas-filled tube.

The Light-versus-current Characteristic.—In addition to the cathode and the cathode-anode characteristics, there is the light-versus-current relationship which is derived from them. This last characteristic is of great practical importance since it is important to know how much current flow is to be expected from a given amount of light. The current flow depends also, of course, on the amount of applied voltage, so that in general a family of curves is needed, as shown in Figs. 99 and 100. These curves show the amount of current flow through the tube

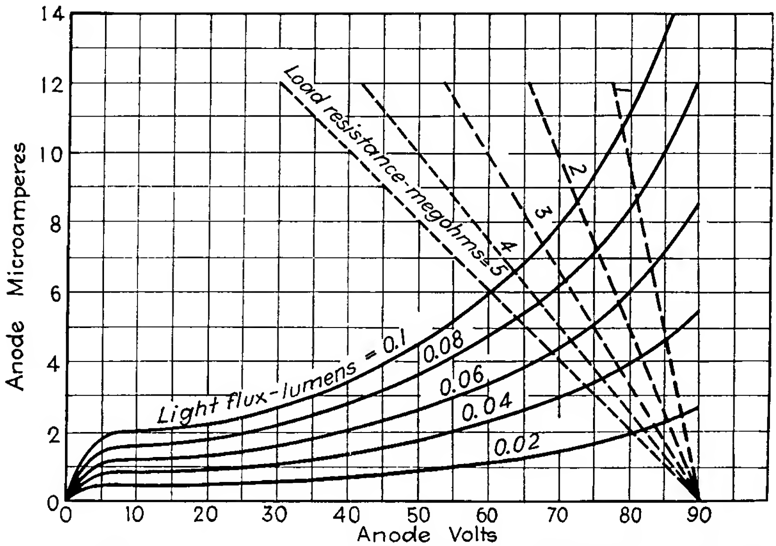


FIG. 98.—Anode current-voltage characteristics of a gas-filled phototube (Type 918).

plotted against the incident light, for several different values of applied voltage. In the case of the vacuum tube (Fig. 99), the relationship between light and current is linear, *i.e.*, the curves are straight lines. This follows from the fact that the current flow is made up purely of photoelectrically emitted electrons, the number of which increases in direct proportion to the amount of incident light. In the case of gas-filled tubes, (Fig. 100), however, the linear relationship no longer exists, since in this case the electrons participating in the current flow come from several sources, some being emitted photoelectrically, some by secondary emission, and some from the other effects

due to the presence of positive ions in the gas. It will be noticed that the current values are some five times as great as in the vacuum tubes.

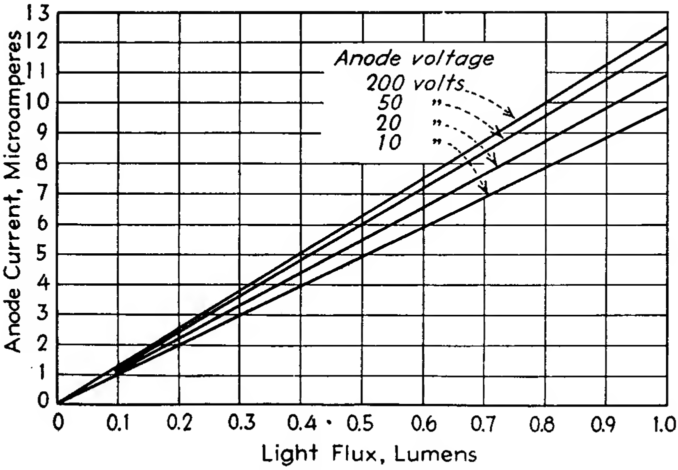


FIG. 99.—Current-versus-light curves of a vacuum phototube (Type PJ-22).

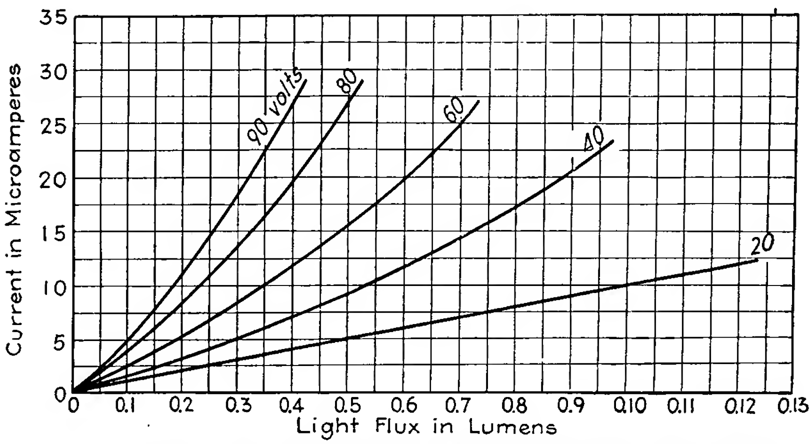


FIG. 100.—Current-versus-light curves of a gas-filled phototube (Type WL735).

The range of incident light considered is from 0 to 1.0 lumen. Usually not more than 2 or 3 lumens of light should be allowed to fall on the cathode if it is to retain full sensitivity, at least not for long periods, and in general smaller amounts of light than this are actually used.

The Mechanism of the Gas-filled Phototube.—The use of gas in a phototube results in current flow several times greater than that obtainable from the same tube pumped free of gas. The principal effect of the gas is to secure a higher value of emission from the cathode than the incident light is capable of extracting. When the operating voltage is applied to the anode (usually not more than 100 volts), the motion of the photoemitted electrons becomes vigorous enough to ionize the gas. The degree of ionization is small, not enough to induce a self-sustained discharge, but it results nevertheless in the production of a large number of positive ions. These ions reduce the negative space charge, but since the density of electrons is not great in the first place the negative space charge is not a limiting factor except at very low applied anode voltages. The ions have another more important effect: they congregate near the negative cathode and there bombard it more or less energetically. The bombardment gives rise to a certain amount of secondary emission which combines with the original photoelectric current. Since the work function of the cathode is low, the tendency to produce secondary electrons is great and the current increase from this source may be considerable.

One other effect is of interest: the ionization process within the body of the gas gives rise to the production of new free electrons which are in part attracted to the anode. This removal of the negative charge from the gas leaves an excess of positive ions. If these positive ions can actually penetrate the cathode surface on bombarding it, they may recombine with an electron within the cathode surface, thereby restoring the charge balance in the external circuit, without the necessity of the electron being emitted at all. It is thus possible that part of the increased current in a gas-filled tube is a true ionic current to the cathode, but there is some question on this point, since it presumably requires more work for an ion to penetrate the cathode than for an electron to emerge from it. In any event the net effect is an increased emission from the cathode and a consequently higher cathode-to-anode current.

As the light on the cathode is increased, the degree of ionization increases, so that both the photoelectric current and the secondary electron current are augmented, the total current increasing more than proportionately to the increase in light.

If the light level is increased far enough, the degree of ionization may increase to the regeneration point, and a self-sustained discharge will begin. It is important, therefore, to protect gas-filled phototubes from both excessive applied voltage and excessive incident light since either, or both acting together, can induce electron regeneration.

Since the current in a gas-filled tube depends upon ionization to a considerable degree, it is necessary to consider the ionization and deionization times concerned, if the tube is to be operated at high frequency, *i.e.*, with current interruptions of many thousands per second. In general the deionization time in a gas-filled tube is very short, since the gas pressure is low and the degree of ionization small. Consequently gas-filled phototubes may be operating at frequencies as high as 10,000 cycles per second without serious loss of output (not more than 10 per cent), owing to deionization lag.

Most of the gas-filled tubes use the caesium-oxide type of cathode, since they are designed for high current output, which the caesium surface is fitted to supply. The effect of gas amplification occurs in the same manner with other cathode materials, but to a lesser degree since they are not so susceptible to secondary emission as caesium.

The term "gas ratio" or "gas amplification constant" is used to express the ratio between the photocurrent with the gas content and that under the same circumstances in vacuum. The values of this ratio vary from 3 to 10 times.

Construction Features of Phototubes.—Several commonly used structural arrangements used in commercial phototubes are shown in Fig. 101. The cathode in one common style (Fig. 101*A*, *B*, *C*.) is a hemicylindrical sheet, with the anode parallel to its axis. The anode is a straight wire of small cross section, so that it will not shield the cathode surface from light. The cathode surface is large, usually a square inch or more, to collect as much light as possible and to provide as large a current as possible from a given light-beam cross section. Since the phototube current is so small, it is important that the insulation between cathode and anode be high, else leakage currents comparable with the photoelectric current may flow between them. These leakage currents mask the desired phototube current. One commonly used construction to reduce the possibility of

leakage is the use of a terminal through the top of the glass envelope. The glass itself is then used as the insulator between the cathode and anode. For extremely exacting work (when low light levels and correspondingly small photocurrents must be preserved), an additional precaution may be taken by coating the glass with a wax to exclude moisture.

A very large cathode surface may be obtained by the use of a spherical glass envelope, on the inside of which is sputtered the

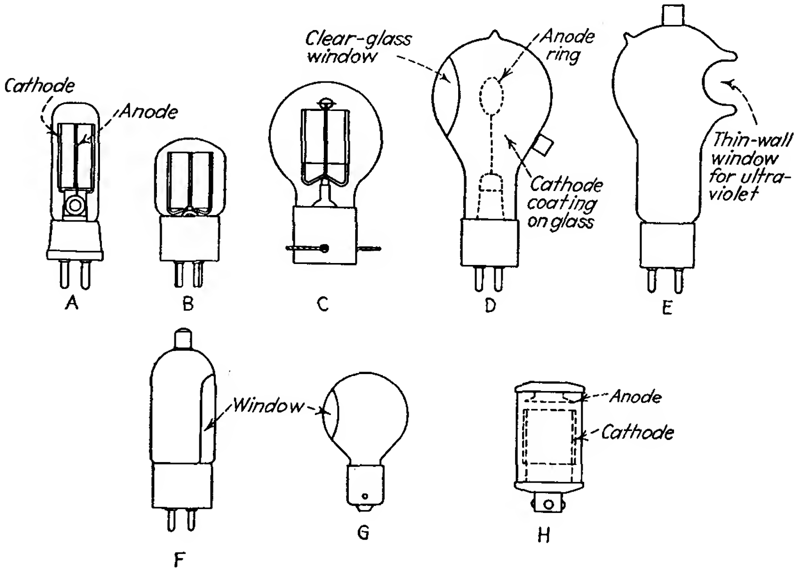


FIG. 101.—Construction of typical phototubes: *A*, *B*, and *C*, hemicylindrical cathodes, wire anodes; *D*, *F*, and *G*, cathodes coated on interior of envelope; *E*, re-entrant-window type; *H*, cartridge type.

cathode material (Fig. 101*D*, *F*, *G*). A small opening ("window") in the side of the tube admits the light, while a ring-shaped anode in the center of the sphere collects the photocurrent. The use of the spherical shape results in multiple reflections of the light within the tube, so that one ray of light will impinge on the cathode surface many times, thus increasing the total number of electrons released by it.

Still another form of envelope is that containing a "re-entrant window" (Fig. 101*E*), a very thin section of glass shaped to withstand the vacuum pressure. This type of window is used to admit light normally not transmitted through the ordinary

thicknesses of glass, particularly in phototubes sensitive to the ultraviolet region.

37. Calculation of Phototube Performance from Characteristics.—The application of phototubes and the calculation of their performance in given circumstances involve two units in addition to the phototube itself: the source of light, its intensity and color composition, and the amplifier which converts the feeble phototube current into a current capable of actuating the desired device. The relationship between the three units is shown symbolically in Fig. 102. The light source is treated for convenience as a point source, but any source whose candlepower distribution is known can be used. If the source is colored, or if a filter is employed, the candlepower and the frequency range of the colored light must both be known. When the candlepower of the point source is known, the number of lumens falling on the phototube cathode can be calculated as from the relationship

$$L = \frac{C \times A}{d^2}, \tag{47}$$

where L is the number of lumens falling on an area A sq. cm. when the area is separated d cm. from a point source of C candlepower. This equation rests on the assumption that the light from the source is sent out equally in straight lines in all directions, and follows from the fact that the total number of lumens radiated is numerically equal to the candlepower of the source. If the source is not a point source, then usually the candlepower in a plane, *e.g.*, the horizontal, is given, which value may be used, provided that the phototube cathode is situated in that plane.

When the number of lumens L falling on the cathode is known, the resulting photoelectric emission from the cathode may be calculated from the sensitivity S of the cathode (in microamperes per lumen), expressed for the particular color composition of the light present. If the voltage applied to the phototube is large enough to produce saturation, then all of the emitted current will flow to the anode. In most practical cases this saturation condition applies, but in the case of gas-filled tubes, the current flow can best be determined by reference to the voltage-current characteristic (page 166).

The amplifier connection is commonly made as indicated in the figure, *i.e.*, by passing the phototube current through a coupling resistor connected in the grid circuit. The passage of the current produces a voltage drop numerically equal to the value of the current (in amperes) times the value of the resistance in ohms. This voltage drop is inserted in series with the cathode and grid of the amplifying tube, and the presence of this voltage alters the current flowing in the plate circuit. To determine the amount of plate-current change thereby produced in the amplifier it is necessary to know the mutual-conductance value of the amplifier tube at the operating current and voltage, and to evaluate also the effect of the load resistance on the plate current. This latter calculation depends on the amplification

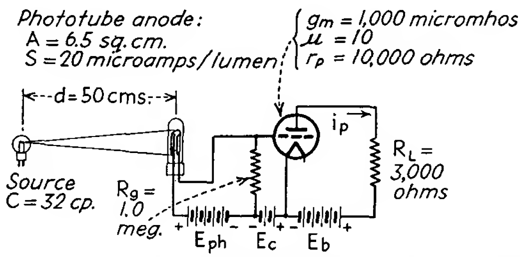


FIG. 102.—Relationship of lamp, phototube and amplifier (numbers refer to Prob. 1).

constant and plate resistance of the tube. The complete solution of the problem involves therefore a thoroughgoing use of both the light-flux relationships and the fundamental amplifier theory discussed on pages 114 to 116. The following problem will illustrate the procedure in a typical case (see Fig. 102):

Problem 1. A 32-candlepower head lamp is situated 50 cm. from a vacuum phototube whose sensitivity to the lamp's light is 20 μa. per lumen. The effective area of the cathode is 6.5 sq. cm. The coupling resistor is 1.0 megohm, the amplifier tube has an amplification factor of 10, a dynamic plate resistance of 10,000 ohms, and a mutual conductance of 1000 micromhos. The load resistance in the amplifier circuit is 3000 ohms. Find the change in amplifier plate current when the lamp is turned on.

Given:

$C = 32$ cp.

$A = 6.5$ sq. cm.

$d = 50$ cm.

$R_g = 1.0$ megohm.

$\mu = 10.$

$r_p = 10{,}000 \text{ ohms} = \dfrac{\mu}{g_m}.$

$g_m = 1000$ micromhos.

$R_L = 3000$ ohms.

$S = 20$ μa. per lumen.

To find: i_p.

The number of lumens L falling on the cathode (assuming equal distribution of light in all directions from the lamp, no lenses or reflectors being used) is:

$$L = \frac{C \times A}{d^2} = \frac{32 \times 6.5}{50^2} = 8.3 \times 10^{-2} \text{ lumens.}$$

If the battery voltage produces saturation-current flow, as is assumed, then the phototube current due to this illumination is

$$i_{ph} = L \times S = 8.3 \times 10^{-2} \times 20 = 1.6 \ \mu\text{a.}$$

This current flowing through the coupling resistor produces a voltage drop.

$$e_g = i_{ph} \times R_g = 1.6 \times 10^{-6} \times 1 \times 10^6 = 1.6 \text{ volts.}$$

With this change in voltage on the grid, the change in voltage across the load resistor is

$$e_L = \frac{e_g(\mu \times R_L)}{r_p + R_L} \text{ [Equation (43), page 115]}$$
$$= \frac{1.6(10 \times 3000)}{10{,}000 + 3000} = 3.7 \text{ volts.}$$

This change in voltage across a 3000-ohm load resistance produces a current change in the resistor (the current change in the plate circuit) of

$$i_p = \frac{e_L}{R_L} = \frac{3.7}{3000} = 1.23 \times 10^{-3} \text{ amp.}$$
$$= 1.23 \text{ ma.}$$

Note that the amplifier has converted a phototube current change of 1.6 μa. into a plate current change of 1.23 ma., a current amplification of nearly 1000 times. This current could be used readily to operate a comparatively rugged, simple, and quick-acting relay, which would then respond to any interruption of the light beam between lamp and phototube.

It is to be noted that the use of the light in the above example is exceedingly wasteful, since only 0.083/32 = 0.0026 of the total light radiated by the lamp actually hits the cathode. By a suitable reflecting and lens system, more than half of the total

available light might be caused to enter the cell. The change in amplifier plate current would then be correspondingly greater were it not for the fact that the currents and voltages would then be so large as to be outside the operating ranges of both phototube and amplifier tube.

38. Photovoltaic Cells.—The photoemissive phototubes just discussed, like all other electron tubes, require a source of voltage to force the electrons to move from the cathode to the anode. Also, since the output current is very small, an auxiliary amplifier is needed. These external-apparatus requirements are overcome in the *photovoltaic cell*, a completely self-contained photosensitive device. The photovoltaic effect, on which the operation of this type of cell depends, is the direct conversion of light energy into electric energy. The process, in the present state of the art, is highly inefficient, only a fraction of a per cent of the light energy being converted.

As previously stated, the photovoltaic effect occurs in solids or liquids and hence is not an electronic process in the restricted sense we have adopted. But in a broader sense it is electronic, since it results from the liberation of electrons at the boundary of a substance under the action of the incident light. The effect is usually associated with a rectifying action, *i.e.*, it occurs at the interface between two materials across which current is carried more easily in one direction than in the other. When this interface is illuminated, electrons are freed and are urged, by a process not very well understood, to one side of the interface where they are trapped by a "blocking layer." That this trapping action is associated with the rectifying action of the interface seems certain, but it apparently acts in the opposite direction to that of the rectifying action. In any event the trapped electrons set up a difference of potential between the two sides of the interface. This voltage difference is discharged slowly across the interface, but if an external circuit is connected, the electrons will prefer to flow through this circuit if its resistance is small compared with that of the interface.

This dependence of the current flow on the use of a low-resistance external circuit, is a limitation on the useful application of the photovoltaic cell. One result is that it is difficult to obtain an appreciable voltage from the cell, since the current flows through a low-resistance circuit. Amplification of the

output of a photovoltaic cell by conventional voltage-operated electron-tube amplifiers is thus not practical. The current available from the cell is large enough to operate many devices directly without amplification, however, so the cell finds many applications. It forms a very convenient illumination meter when connected with a portable, rugged, low-resistance micro-ammeter, and it can be made to operate low-current relays. Since the power produced by the cell is small, the relays are in general of the low-torque, slow-acting variety, and are consequently of restricted use where rapid dynamic response is required.

Two types of photovoltaic cell now in wide use are the cuprous oxide type and the selenium-iron type, the names referring to

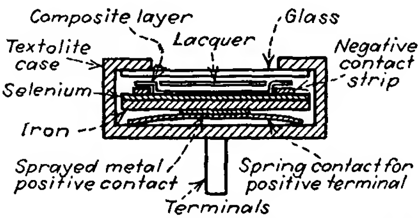


FIG. 103.—Construction of selenium-on-iron self-generating (photovoltaic) cell.

the type of rectifying substance used to obtain the current. Cuprous oxide, formed by heating pure copper to 1000°C. in air, has long been known for its rectifying properties but its photovoltaic action was not discovered until 1926, and its commercial application was first undertaken in 1930. The commercial form of the cell consists of a disk of copper on which is the cuprous oxide layer. A very thin metallic coating is sputtered on the oxide. The copper forms one terminal of the cell, the metallic layer the other. The active interface is the boundary between the copper and the oxide.

The selenium-on-iron cells (Fig. 103), which also possess rectifying properties, consist of a disk of iron on which is coated a layer of selenium, with which is mixed a trace of impurity in the form of silver. On the selenium is deposited a thin sputtered layer which acts as one terminal of the cell.

Characteristics of Photovoltaic Cells.—The characteristics of photovoltaic cells express the current output, the voltage output, and the spectral response. The current sensitivity has values

from 100 to 500 μa. per lumen, but the sensitivity changes with the degree of illumination. The current output depends, of course, on the amount of resistance connected in the external circuit, and this fact is taken into account in the light-versus-current curves for each cell (see Figs. 104 and 105). The

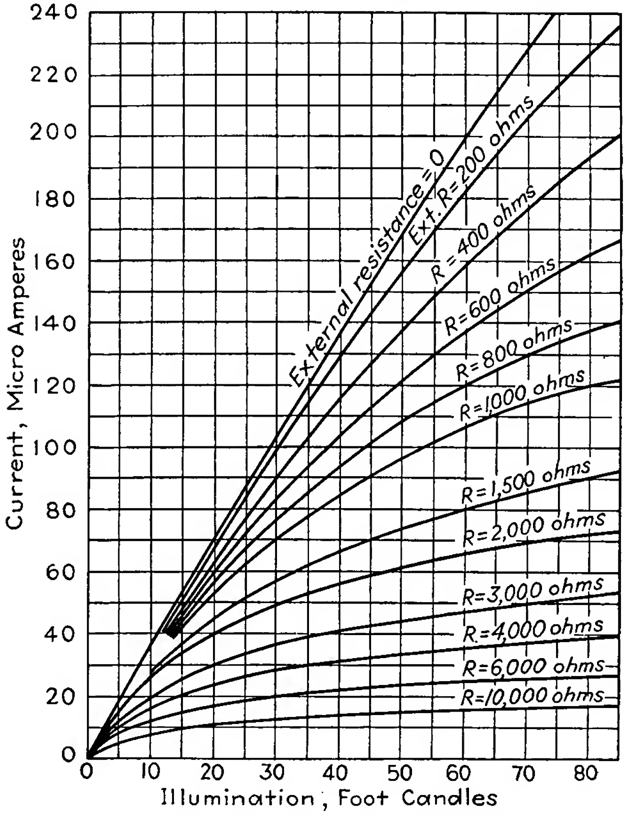


FIG. 104.—Current versus illumination of a typical photovoltaic cell. Note that high output currents are obtainable only in low-resistance circuits.

current is not strictly proportional to light owing to the action of space charge at the interface. The voltage output reaches only 0.5 volt, even in the best cells. The spectral response depends on the materials used in the cell and on its construction. If the light is caused to go through the oxide or selenium layer before hitting the interface (back-effect cell), the response is best in the red region, whereas if it passes through the interface

first (front-effect cell, not commonly used commercially), the response is greatest in the blue region. The spectral response of a typical photovoltaic cell is shown in Fig. 106. The selenium-iron type of cell, used in conjunction with a newly developed filter, has a response very similar to that of the human eye, which makes the combination especially useful in applications in which the cell must replace the eye.

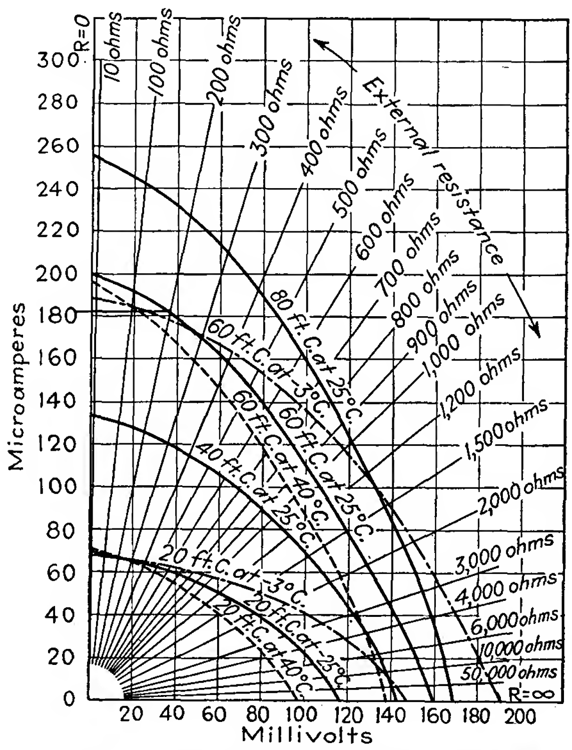


FIG. 105.—Complete photovoltaic cell characteristics, relating current and voltage output to light, circuit resistance, and temperature.

A simple type of photovoltaic cell not used to any extent commercially is the electrolytic type, in which an interface forms between a liquid and a metallic electrode.

39. Photoconductive Cells.—The oldest form of electrical photosensitive device is the selenium cell, a form of photoconductive cell. Selenium is not the only substance which displays the effect, since some 15 or more natural minerals have

been found with photoconductive properties, but it is the only material used to any extent. Selenium is a semiconductor, that is, a material on the border line between the conductors and insulators. When it is illuminated, its electrical resistance decreases, and the current flow for a given applied voltage

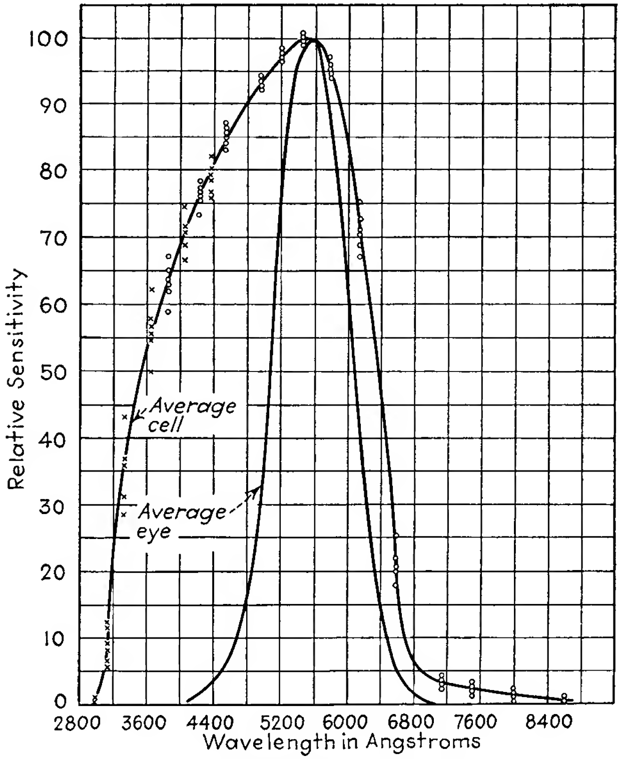


FIG. 106.—Photovoltaic cell spectral response, compared with the response of the human eye.

increases, presumably because of the liberation of additional unbound electrons within the material.

The amount of current change available for a given change in light is higher in the selenium cell than in the photovoltaic cells or in phototubes, but its characteristics otherwise are not desirable. In the first place, the amount of current increase is not directly proportional to the increase in light. This results from the fact that the freed electrons within the selenium set up

a space charge which increases as the light increases. This space charge inhibits the current flow; finally producing a saturation level, beyond which the current will not increase with increasing light. This same effect makes the response of photoconductive cells rather sluggish. Finally ordinary selenium cells have a

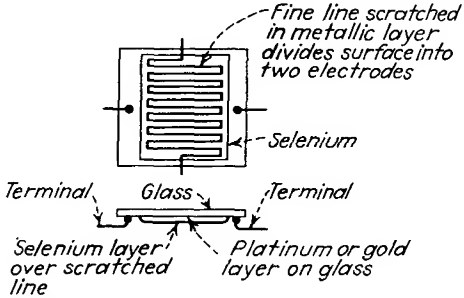


FIG. 107.—Construction of a typical selenium photoconductive cell.

rather large temperature coefficient, which limits their application in exposed locations. Modern forms of selenium cell, manufactured and operated in vacuum, have a much lower temperature coefficient.

Since the selenium is not a perfect insulator, a certain amount of current flows through it whenever voltage is applied to it, even though no light falls on its surface. This current is called the *dark current*, and it is often troublesome since the control provided by the cell is a marginal control only.

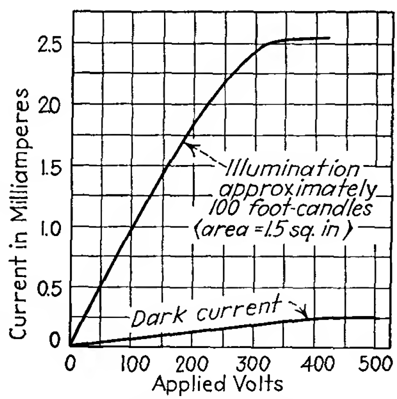


FIG. 108.—Current output of a selenium cell versus applied voltage.

A typical current-voltage curve of a selenium cell is shown in Fig. 108. The current depends, as in the photoemissive cell, on both illumination and applied voltage. The need of an external battery, the general instability of its characteristics, the time lag, and the large dark current present in the selenium cell have restricted its use.

Problems

1. A 32-candlepower headlight lamp is situated 1.5 ft. from a phototube whose sensitivity to the light produced is 50 μa. per lumen. The phototube is coupled through a 2.0-megohm resistor to the grid of a triode amplifier tube whose amplification factor is 9 and whose mutual conductance is 1100 micromhos. When the lamp is turned on, what change in plate current is recorded, (*a*) when the plate load is a milliammeter of negligible resistance, and (*b*) when the plate load is a relay whose coil has a resistance of 2000 ohms? The cathode area of the phototube is 0.9 sq. in. No lenses or reflectors are used, and the lamp is considered as simple point source.

2. If the lamp in Prob. 1 is replaced by a 0.5-candlepower flashlight lamp, what is the plate-current change in cases (*a*) and (*b*)? If the flashlight lamp is brought within 3 in. of the phototube cathode what are the current values?

3. A lens of 1 in. diameter situated 5 in. from a 0.5-cp. lamp directs a parallel beam of light to the phototube and amplifier described in Prob. 1. What plate-current change occurs when the lamp is turned on, in cases (*a*) and (*b*)? (HINT: Assume all the light striking the lens enters the phototube.)

4. A certain caesium-oxide photosensitive cathode has a sensitivity of 20μa. per lumen when exposed to the light of a tungsten 60-watt lamp filament operated at 2870°K. What is the sensitivity of the cathode to the same source fitted with an ideal filter which passes all the visible wave lengths, from 4000 to 7000Å inclusive, but blocks out all other wave lengths. The relative energies in the wave lengths from 3000 to 11,000Å given off by the tungsten lamp at 2870°K. are given in the table below. [HINT: Multiply each of these energy values individually by the corresponding ordinate in the caesium—cathode spectral-response curve (Fig. 94). Compare the sum of all these products with the sum of those between 4000 and 7000Å.]

RELATIVE ENERGY VALUES

Wave length	Relative energy	Wave length	Relative energy	Wave length	Relative energy
3000	0	6000	45	9000	100
3500	1	6500	60	9500	100
4000	5	7000	70	10000	100
4500	10	7500	80	10500	100
5000	20	8000	90	11000	95
5500	35	8500	95		

5. Repeat Prob. 4 for an ideal filter passing light, without loss, between 7000 and 10,000 Å. inclusive.

6. A photovoltaic cell having the current, voltage and light characteristics of Fig. 105 receives 60 ft.c. of light at 25°C. What is the power output

of the cell when connected to a 200-ohm resistor? If 1 lumen of light is equivalent to 0.002 watt (for the color used), what is the efficiency of energy transfer from light to resistor? The cell area exposed is 2 sq. in., (1 lumen per square foot = 1 foot-candle).

Bibliography

HUGHES and DUBRIDGE: "Photoelectric Phenomena," McGraw-Hill Book Company, Inc., New York, 1932.

CAMPBELL and RITCHIE: "Photoelectric Cells," Isaac Pitman & Sons, 1929.

WALKER and LANCE: "Photoelectric Cells and Their Application," Isaac Pitman & Sons, 1935.

ZWORYKIN and WILSON: "Photocells and Their Application," John Wiley & Sons, Inc., New York, 1934.

KOLLER, L. R.: "Physics and Electron Tubes," McGraw-Hill Book Company, Inc., New York, 1937, Chaps. XII, XIII, XIV.

MACARTHUR, E. D.: "Electronics and Electron Tubes," John Wiley & Sons, Inc., New York, 1936, Chap. IV.

DOW, W. G.: "Fundamentals of Engineering Electronics," John Wiley & Sons, Inc., New York, 1937, Chap. XVIII.

HENNEY, KEITH: "Electron Tubes in Industry," McGraw-Hill Book Company, Inc., New York, 1937, Chap. V.

CHAPTER IX

ELECTRONIC SOURCES OF LIGHT

Introduction.—As pointed out in Chap. V, the passage of electrons through a gas or vapor is accompanied by interchanges of energy between the electrons and the gas or vapor molecules. These interchanges may be mechanical energy, electrical energy, or light energy. The present chapter is concerned with the last-named type of energy transformation, the electronic production of light from electricity. Gas- and vapor-filled lamps, the practical devices which make use of this process, are at present the subject of a great amount of research in commercial lamp companies. While these lamps are of small commercial importance, compared with incandescent-filament lamps, they hold promise of providing cheaper light for longer periods. One of the newest members of the vapor-lamp family, the sodium lamp, is the most efficient source of artificial light now known. In one field, outdoor advertising signs, the electronic form of light in the form of neon-, helium-, and mercury-vapor-filled tubes has all but supplanted the incandescent bulb. The electronic production of light is thus a subject of growing importance in practical engineering.

40. The Mechanism of Light Production in Gas Discharges.—The light output of any source can be described in terms of its spectral distribution, *i.e.*, the relative intensity of the component colors it contains. When the light is produced by an incandescent solid, such as the filament of a lamp, it is found that all colors in the visible range are contained in it, that is, its spectrum is continuous. The light produced by a gas discharge, on the other hand, is characterized by a discontinuous spectrum, and its spectrum is made up of discrete lines or bands of color. Several examples are given in Fig. 109. If the gas is monatomic (one atom per molecule, as it is in all the gases and vapors commonly used), the spectrum is made up of lines; if it is polyatomic (more than one atom per molecule, carbon dioxide being

the only important example), the spectrum is made up of bands.

The presence of the spectrum lines was accounted for by Bohr, who proposed in 1913 that the light from each molecule is radiated when the molecule changes from one energy state to another (*cf.* Sec. 20, page 72). He suggested that the change in energy might correspond to the removal of an electron from one orbit to another.

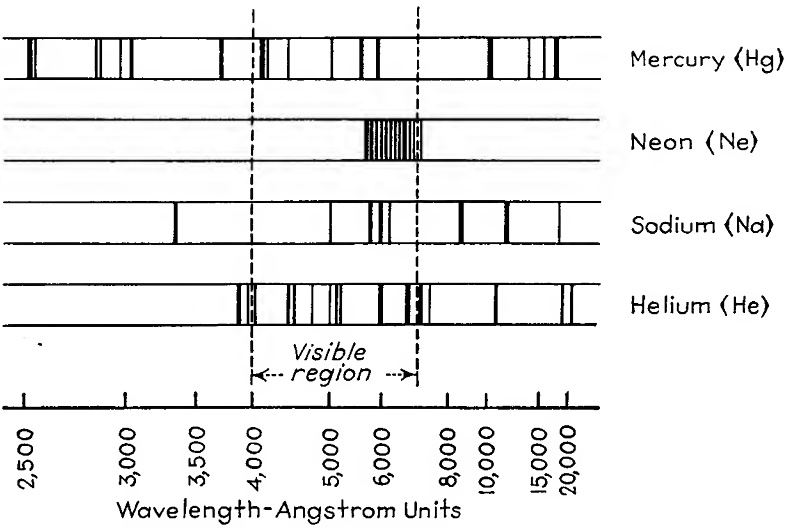


FIG. 109.—Line spectra of gases used in electronic lamps.

The Bohr theory states that if the atom changes from an initial energy of W_i ergs to a final energy of W_f ergs, then the frequency f of the emitted radiation is given by the relation

$$f = \frac{W_i - W_f}{h} \text{ waves per second,} \tag{48}$$

where h is Planck's constant, 6.54×10^{-27} erg-sec. This frequency specifies the color of the light (red being a low-frequency, violet a high-frequency radiation), but it is more customary to speak of the color of light in terms of its wave length, measured in angstrom units, 1Å. being 10^{-8} cm.

The wave length and frequency of light are related to the speed at which it travels, by the following reasoning: If a source of light is sending out f waves per second, and if each wave is

λ cm. long, then the total length of the beam at the end of 1 sec. must be $f\lambda$ cm. and the speed c of the light must $f\lambda$ cm./sec., therefore

$$f\lambda = c \quad \text{cm./sec.} \tag{49}$$

The measured speed of light, c, is about 3×10^{10} cm./sec. Hence, if the frequency of the light is known, its wave length can be computed.

It is convenient to express the energy levels of the radiating atom not in ergs but in equivalent electron volts (1 electron volt is 1.6×10^{-12} erg). Therefore, substituting in Eq. (48), we obtain for the wave length λ of the emitted light, in centimeters,

$$\lambda = \frac{c}{f} = \frac{ch}{W_i - W_f} = \frac{ch}{1.6 \times 10^{-12}E} = \frac{3 \times 10^{10} \times 6.54 \times 10^{-27}}{1.6 \times 10^{-12}E} \text{ cm.}$$

which, expressed in angstrom units, is

$$\lambda = \frac{12{,}336}{E} \tag{50}$$

This is a very convenient equation, since it gives the wave length of the emitted light directly in terms of the energy change E, expressed in volts, which the atom undergoes. It is plotted in Fig. 110.

The limiting wave lengths of visible light are approximately 4000Å. at the violet end and 7000Å. at the red. The corresponding voltage values, computed from Eq. (50) are 3.09 volts for the violet and 1.76 volts in the red. Hence if the energy change in the atom is less than 1.76 volts, the emitted light lies in the infrared region; if it is greater than 3.08 volts, it lies in the ultraviolet. Thus the energy changes which can produce visible light in a gas discharge are restricted to a very narrow range. This fact has an important bearing on the relative efficiency of light production of different gases and vapors.

Energy Transfers from Electron to Atom.—The changes in energy referred to above may come from any source, but in gas discharges they are produced primarily by the impact of a free electron on an atom. The quantum condition (page 74) restricts the amount of energy which the atom may accept from

the electron. If the electron has been accelerated through a voltage drop of less than 2 volts, its collision with the atom results in a simple elastic rebound, and the atom receives virtually no energy at all. But if the energy of the electron is 2 volts or higher, its energy may be passed completely to the atom, depending on the type of gas or vapor in question. In sodium vapor

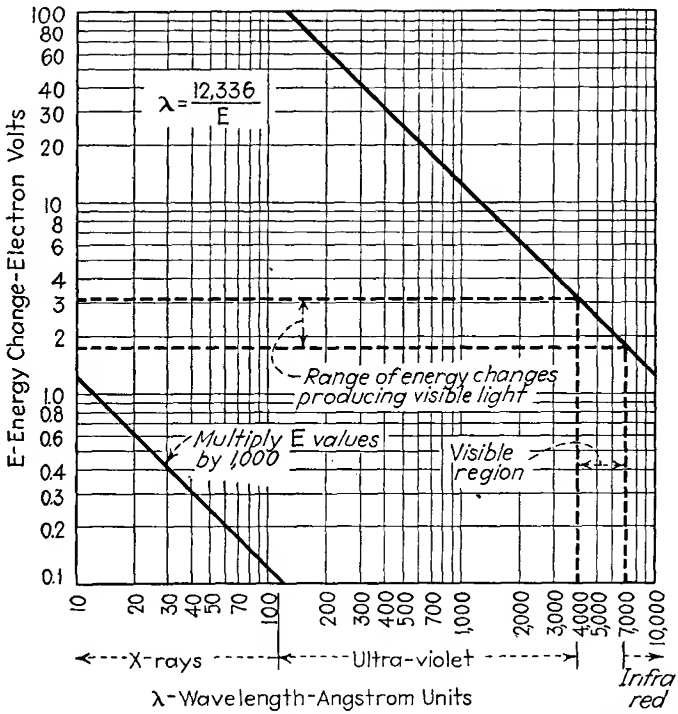


FIG. 110.—Wave length of light produced by electronic energy changes [plot of Eq. (50)]. Visible light results from energy changes having values between 1.76 and 3.09 volts.

this complete energy transference occurs when the electronic energy is 2.1 volts, in mercury vapor at 4.86 volts, in neon at 16.6 volts, in helium at 20.9 volts.

When the atom has accepted this energy from the electron, it retains it for a very short but still appreciable length of time, approximately 10^{-7} to 10^{-8} sec., depending on the energy level itself. At the end of this time, the atom suddenly reverts to its normal energy level, and the extra energy it has absorbed

from the impacting electron is released as radiant energy. The wave length of this energy is calculable [Equation (50)] from the voltage difference in which the energy change is expressed. According to the above voltage values, the light radiated by sodium is about 5890Å., mercury, 2540Å., neon, 740Å., and helium 592Å. Of all these radiations only one, that of sodium, resides in the visible region between 4000 and 7000Å. All the others are in the ultraviolet.

The fact that all of the gases and vapors cited actually do produce visible light seems to be in direct contradiction to this

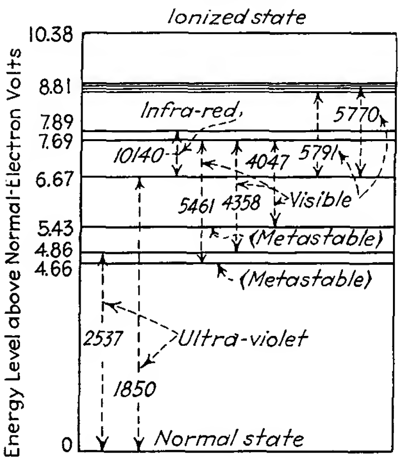


FIG. 111.—Energy levels of the mercury atom, showing the production of visible light by changes between excited levels.

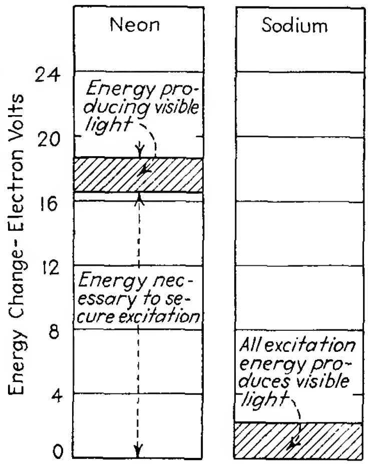


FIG. 112.—Energy levels of neon and sodium compared.

statement. But the answer lies in the fact that the atom may have many energy levels higher than those stated. The 2.1-volt level for sodium is the lowest level, but there are others higher than this which can be excited by a high-energy electron. An atom, so excited to a higher energy level, can return to its normal state in a series of jumps, and the energy change involved in each jump may lie within the range which produces visible light. The visible-light output of all the gases and vapors mentioned, except sodium, is thus produced by changes in energy between two "excited" levels, and not between an excited level and the normal level. To produce visible light from neon, for example (see Fig. 112), it is necessary to have an

electronic energy approximately 2 volts higher than its lowest excited level, or, $16.6 + 2 = 18.6$ volts. When the atom accepts this energy, it may jump first to the 16.6-volt level, and then to its normal level. The first jump, being only 2 volts, will produce radiation within the visible region; the second jump will not. Thus only 2 parts in 18.6 of the original energy are used to produce useful (*i.e.*, visible) light. In sodium vapor, on the other hand, a 2.1-volt electron will excite the atom, and the atom, on returning to normal will use this full amount in the production of visible light. It is to be expected, therefore, that the production of light from sodium vapor should be much more efficient than that from neon, and this is indeed the case.

There is a considerable number of energy levels above the first excited level in each atom. In sodium, for example, there are about 15 levels available, and the energy output may result from an energy change between any two of these 15. In mercury (see Fig. 111) and neon, the situation is even more complex. The light output from these gases and vapors is composed, therefore, not of one or two single wave lengths, but of a great many. The line spectrum of mercury contains six important components in the visible region, in the violet, blue, green, and yellow regions. Neon has about 15 component colors (lines) all clustered in the yellow, orange, and red regions. Helium has a component in nearly every portion of the spectrum, which accounts for its whitish color (white light is composed of light of all the spectrum colors).

Absorption and Retransmission of Light.—The energy temporarily absorbed by the gas atom need not be supplied by an impacting electron. It may as well be absorbed from the energy of a light beam. But, just as the atom will accept energy from electrons only in the proper "allowed" amounts, so also will the atom accept energy from light only if the light is of the proper wave length. In particular, atoms will accept energy from light which has the same wave length as the light which the atom will itself produce. Thus we find that if mercury vapor is irradiated with light of 2540Å. wave length, the mercury vapor will absorb the light. But immediately thereafter (within 10^{-8} sec.) it will re-emit light of the same wave length. This light may in turn be absorbed again by near-by mercury atoms, and re-emitted by them. In fact in a gas-filled lamp the light is

continually absorbed, emitted, reabsorbed, re-emitted, and so on. In a sodium vapor lamp 2.5 cm. in radius, for example, it is estimated that the light is absorbed and re-emitted 100,000 times before it reaches the glass envelope and escapes through it. Since each absorption period lasts for 10^{-8} sec., the light actually consumes about a thousandth of a second in traveling across the lamp to the glass surface. This is an extremely long period of time, speaking in terms of the velocity of light and the times usually required for atomic transformations. The phenomenon has, in fact, been called "imprisonment of radiation." The light actually radiated from a gas lamp, which must come, of course, from the gas immediately inside the bulb, is thus not the original electrically produced light, but light which has been passed, bucket-brigade fashion, from the original impacted atom in a succession of many thousand transfers.

41. Characteristics of Gases and Vapors Commonly Used in Lamps.—A very great number of gases are available for use in gas- and vapor-filled lamps, and many of them have been very carefully examined. Compton in 1923 studied the vapors of lithium, potassium, rubidium, caesium, cadmium, zinc, and the rare gases argon, xenon, and krypton. But all of them produced a very feeble light in relation to the electrical energy used, and hence are of doubtful commercial value. Molecules of the common gases nitrogen, oxygen, hydrogen, and carbon dioxide

TABLE IX.—LUMINOUS PROPERTIES OF GASES AND VAPORS COMMONLY USED AS LIGHT SOURCES[1]

Gas or vapor	Predominant color	Ideal luminous efficiency, lumens per watt	Specific luminous efficiency, lumens per watt	Energy-utilization ratio
Sodium	Yellow	475	50–100	0.10 –0.21
Mercury (low pressure)	Blue green	248	15–20	0.06 –0.08
Mercury (high pressure)	Blue green	298	35– 35	0.10 –0.12
Neon	Orange red	198	15–40	0.075–0.20
Helium	Yellow white	...	4–10	
Carbon dioxide	White	...	2–4	
Flaming arc	White	220	27–45	

[1] DUSHMAN, *Gen. Elec. Rev.*, **37**, 262 (1934).

have also been investigated, but carbon dioxide is the only one ever used. In fact, the carbon dioxide-filled Moore tubes, which made their appearance in 1906, were the first gas-filled tubes used for advertising purposes. Carbon dioxide is active chemically, and tends to combine with the metal of the electrodes in the lamp. The Moore tubes contained an automatic magnetically-operated valve for admitting a fresh supply of the gas as it was needed. Carbon dioxide lamps are still used to some extent for color matching, since the gas gives a white light.

The gases and vapors actually used at present are limited to five: the rare gases neon, helium, and argon (the last only in mixtures), and the vapors of mercury and sodium. Neon produces a red-orange light of great intrinsic brilliance and is therefore much used for advertising purposes. Its specific luminous efficiency (a measure of the amount of light energy produced per unit of electrical energy supplied to the gas) is from 15 to 40 lumens per watt, depending on the type of construction and the auxiliary equipment used. Helium gives a yellow-white light; its efficiency is 4 to 10 lumens per watt. Mercury produces 30 to 35 lumens of blue-green light per watt. Sodium, most efficient of all, produces 50 to 100 lumens of yellow light per watt. Carbon dioxide, in comparison, has an efficiency of 2 to 4 lumens per watt.

The rare gases have the advantage that they are completely inert and consequently will not combine chemically with the electrodes in the lamp. Mercury vapor, the pressure of which is automatically maintained by vaporization from the liquid mercury sealed within the tube, has the disadvantage that its pressure varies over wide limits with the temperature of the tube. This limits the operation of the tube, especially at the low temperatures encountered in exposed locations. It will also combine with any impurities present in the tube to produce a black discoloration of the glass, so that great care in cleaning the glass and the electrodes must be exercised in the manufacture of the tubes.

Sodium vapor is very active chemically and will attack ordinary glass, so that special resistant glass must be used for the envelope. At room temperature (25°C.) sodium is a solid, and the vapor pressure from it is so low that no discharge is available from it under any circumstances. Even at operating temperature,

the vapor pressure is only 0.0002 mm., and the discharge can be maintained only in the presence of neon gas which is mixed with the vapor in all practical sodium lamps. The mechanism of this partnership between sodium and neon is described on page 193.

42. The Electrical Action in the Luminous Discharge. Cathodic and Positive Column Forms.—The light produced in a gas discharge is the result of excitation, from the temporary addition of energy to the gas atoms. But the current carried by the discharge depends not on the excitation of the atoms but on their ionization. The problem of the light-producing discharge is thus twofold; (1) excitation for light and (2) ionization for current flow. The electrical action of the discharge is explained primarily in terms of ion production.

When voltage is applied between two electrodes in the gas, the resultant electric field accelerates the free electrons present, imparting energy to them. If the energy picked up is large enough (if the voltage drop per mean free path is high), some of the collisions between electrons and atoms ionize the atoms. The positive ions reduce the space charge within the tube, producing a large current between cathode and anode, and collect near the cathode, there producing a thin (about 0.1 mm. thick) sheath of ions. The region between this sheath and the cathode itself is one of high electric field, so the voltage change just outside the cathode surface is very great. This is shown symbolically in Fig. 83 (page 144). The high field, through the agency of field emission and secondary emission, extracts electrons from the cathode. If the cathode is thermionic, the field necessary to extract the electrons is much less; in fact the electrons may become free without the necessity of any field at all. But in any case the electrons, upon leaving the cathode, find themselves suddenly accelerated to high speed as they pass through the cathode sheath. They leave the sheath with high kinetic energy and collide, sooner or later, with a neutral molecule. If the energy of the electron is high enough, part of its energy is transferred to the molecule, and if the energy exceeds the ionization voltage of the gas, the molecule becomes ionized. Actually the energies of the electrons have a wide range of values. Some have low energy, with no resulting energy transfer. Some have enough energy to excite the molecule and so produce light, but

not enough to remove an electron from the molecule. Finally, some have sufficient energy to ionize the molecule. This last event must occur frequently enough to produce a large number of ions, otherwise electron regeneration will not occur and the discharge will not be self-sustained. But if the applied voltage is high enough in relation to the mean free path of the electrons in the molecules of gas, then ionization occurs, and with it, of course, a large amount of excitation from which the light is produced.

It is important to remember that almost all of the energy which the free electrons pick up is obtained during the passage of the electrons through the ionic sheath at the cathode. When the electrons have passed through this sheath, there is very little remaining voltage drop through which they can gain further energy. Hence the tendency is for the electrons to lose energy in successive impacts with molecules, until finally the energy remaining with the electron is lower than that necessary to bring the molecules to their lowest excited level. This low-energy electron is called an ultimate electron; it collides elastically with the molecules until finally it is collected by the anode and returned to the cathode through the external circuit. It may then be re-emitted from the cathode, gaining new energy as it passes through the cathode sheath.

With these facts in mind, it will be seen that the design of an efficient gas-filled lamp involves the choice of the proper gas or vapor to produce the desired type of light with the greatest efficiency, the design of electrodes, gas pressure, and electrode separation to provide energetic electron impacts for ionization and excitation of the molecules, and finally the provision of conditions such that the gas molecules pass through a large number of energy transitions within the visible-radiation region (1.76 to 3.09 volts). Much progress in gas-discharge lighting has been made by the engineering application of these requirements.

Cathodic Versus Positive-column Types of Discharge.—There are two important forms in which the gas discharge is used in practical gas-filled light sources, the cathodic and positive-column discharges. The cathodic discharge, shown in Fig. 113 is more or less compact, the spacing between the electrodes being small compared with the diameter of the bulb itself. The positive-column discharge (Fig. 114), on the other hand, is

very elongated, having a very large separation between the electrodes compared with the diameter of the tube. In the cathodic discharge, practically all the voltage drop in the tube is confined to the ionic cathode sheath, and there is practically no field acting anywhere else in the tube. The emitted electrons produce ionization first; then, having less energy, they produce simply excitation. In the positive-column type, the major voltage drop is also in the ionic sheath, but there is also a slight voltage drop along the rest of the length of the tube, called the positive column. This voltage drop urges the low-energy electrons toward the anode. As they move along the tube, they gain sufficient energy to cause further excitation of the atoms, and as a result the whole length of the positive column glows. The length of the positive column adapts itself to the length of the tube. In neon advertising signs, it is often 70 ft. long, in a tube $\frac{3}{8}$ in. in diameter.

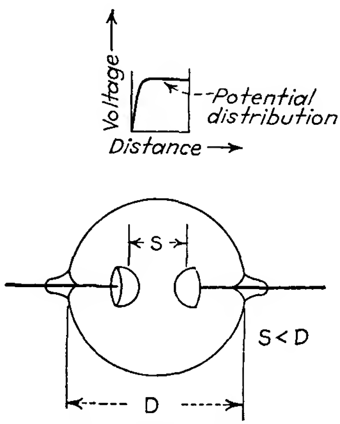


FIG. 113.—Cathodic type gas-discharge lamp.

In both cathodic and positive-column discharges, the cathode is covered with a luminous *cathode glow* (produced, at least in part, by recombination of ions and electrons), which is of little importance from a light-production standpoint. If the separation between the electrode is very small, so that the electrons cannot move far enough to encounter a neutral molecule before

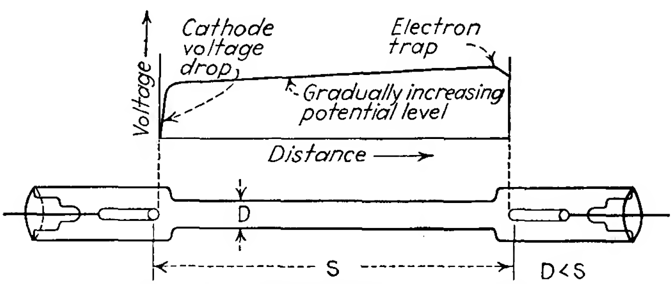


FIG. 114.—Positive-column type of lamp.

they hit the anode, then the cathodic or positive-column discharge disappears and the only source of light is the cathode glow. Such an arrangement is used in the common neon glow lamp.

An important electrical characteristic of any gas-discharge device, and hence of all gas-discharge lamps, is the fact that the current flow through the tube is limited not by the gas-conduction path but by the value of the resistance (or reactance, see Chap. XI) in the external circuit. If this external resistance is small, the current through the lamp tends to be excessive and overheating of the entire structure occurs. To guard against such effects all gas-filled lamps are used in conjunction with a series resistance (or reactance) control which limits the current to the safe operating value.

The Neon-sodium Discharge.—The high efficiency of sodium vapor as a light source has given it outstanding importance in the gas-conduction field. Remarkably enough, sodium by itself is useless as a light source, since its vapor pressure at ordinary operating pressures is so low, and the number of ionizable molecules so few, that ionization does not occur. However, if neon gas, at a pressure of about 1.5 mm., is added to the tube containing the sodium, it performs two important functions: In the first place, the neon provides a path for the start of the discharge. When the voltage is first applied, therefore, sodium lamps glow with the characteristic red-orange color of neon and their luminous output has the efficiency associated with neon gas. As the discharge continues the temperature of the tube gradually increases to the operating point of about 220°C. The vapor pressure of the sodium, increasing with the temperature, has then reached a value of 0.0002 mm., a value still much too small to support ionization by itself. But at this point there forms a partnership between the neon and sodium atoms.

The ratio of pressures indicates that there are about 8000 neon atoms present for every sodium atom. The neon atoms require 16.6 volts energy for excitation, the sodium atoms 2.1 volts. An electron emitted from the cathode, and passing through the cathode sheath of neon ions, gains enough energy to excite the neon atoms. It also has enough energy to excite a sodium atom many times over, but the sodium atoms are so rare that the

electron does not encounter them and hence cannot excite them. It does encounter the neon atoms, exciting some, ionizing others. The electron bounds back and forth between the neon atoms, losing energy at each impact. This succession of impacts keeps the original electron in circulation for a long time, and hence increases greatly its chances of striking a sodium atom. When the electron energy has fallen below 16.6 volts, it is powerless to excite the neon directly (except by a step-by-step cumulative process), but it still has plenty of energy to ionize or excite the sodium, providing it meets a sodium atom. During its long

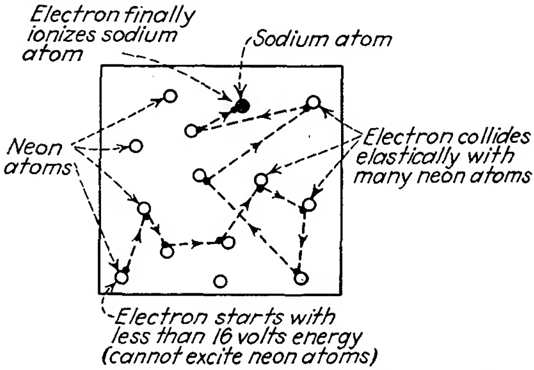


FIG. 115.—Mechanism of the neon-sodium discharge.

succession of elastic collisions with the neon atoms, the electron finally does meet a sodium atom and excitation or ionization at once occurs. The excitation produces a great deal of light, since the sodium atom is very efficient. In fact, when the sodium discharge begins, its light overshadows the neon and the light changes from red-orange to brilliant yellow. The necessity for continual ionization and excitation of the neon still exists, however.

43. Practical Electronic Light Sources.—Of the many forms which gas-discharge lamps have taken the most important commercially are the Cooper-Hewitt lamp, the neon-, helium-, and mercury-vapor-filled tubular units used for advertising and display purposes, all of which are positive-column types, the various forms of low- and high-pressure mercury-vapor lamps and the sodium-neon lamp.

The Cooper-Hewitt lamp makes use of a mercury-pool cathode; the cathode spot is started by the application of a high starting

voltage to an external starting electrode. A typical unit produces about 5000 lumens at 12 or 13 lumens per watt. While the oldest form of gas-discharge lamp, they are still extensively used for factory lighting, and in certain types of photographic processing.

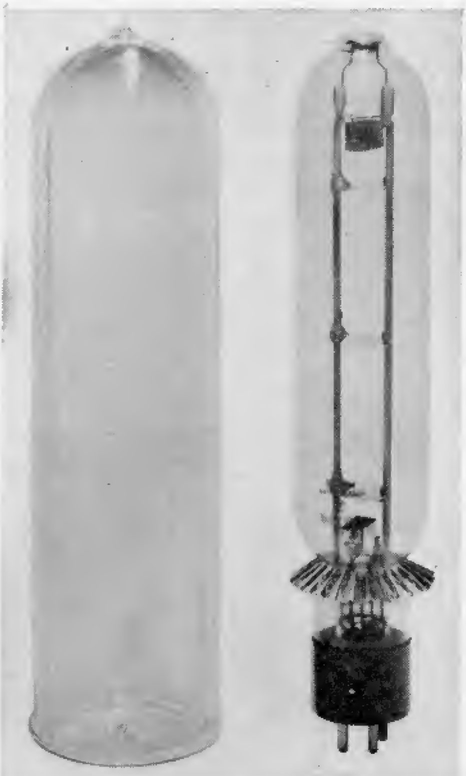


FIG. 116.—Typical sodium-discharge lamp. The envelope at the left is a vacuum flask which slips over the lamp, conserving heat and protecting it from temperature changes.

The tubular units commonly called "neon signs" contain cold cathodes, usually not treated to enhance emission, and are filled with rare gas. The neon-filled (red-orange) tubes contain the gas at about 10 mm. pressure; helium (yellow) is commonly used at 3 mm. pressure. The mercury-filled tubes (blue) contain liquid mercury and rare gas mixed, usually argon or argon and neon at about 8 mm. pressure. The rare gas produces

very little light, compared with that from the mercury vapor, but it does provide a heating effect which supports the mercury-vapor pressure. The tubes vary in length from 1 to 70 ft. in length, depending on the use. The shorter tubes are fed from 2000-volt transformers, the longer from 15,000-volt units. The operating current, 18 to 60 ma., is maintained by the design of the transformer which will not pass more than the rated amount, even when short-circuited, owing to the high leakage reactance incorporated in its core design. Colored glass is often used for the envelope of these tube lights, the green color being produced by mercury-argon in a yellow envelope. By combining the primary color of the radiation with the selective properties of colored glass it is possible to produce a very wide range of colors. The luminous efficiency depends on the diameter of the tubing, the type of gas-filling, and on the absorption properties of the glass; it ranges from 2 to 13 lumens per watt.

Figures 116 and 117 show typical forms of sodium- and mercury-vapor lamps. In the sodium tube a vacuum envelope

Table X.—Gases, Vapors, and Colored-tubing Combinations Used in Luminous Tubes

Color produced	Gas or vapor used	Color of tubing	Pressure of rare gas, mm.
Red	Neon	Clear	10–20
Dark red	Neon	Soft red	10–20
Soft red	Neon	Opal	10–20
Light blue	Mercury, argon, neon	Clear	8
Dark blue	Mercury, argon, neon	Soft blue	8
Light green	Mercury, argon, neon	Canary	8
Medium green	Mercury, argon, neon	Yellow	8
Dark green	Mercury, argon, neon	Amber	8
White	Helium	Clear	3
Soft white	Helium	Opal	3
Gold	Helium	Canary	3
Orange	Neon	Yellow	10–20
Red lavender	Neon	Purple	10–20

is used, to provide heat insulation. One of the most promising developments is the high-pressure mercury lamp (Fig. 117) which produces an extremely efficient and brilliant light of white-blue color. The spectral composition of this light is quite different from that of the low-pressure mercury lamps. Inclosed

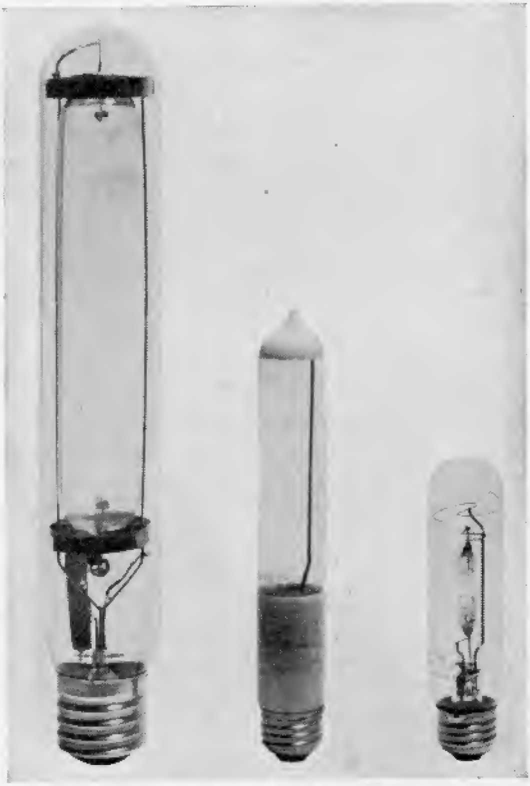


FIG. 117.—Types of mercury-discharge lamps. The "capsule" type at the right operates at very high vapor pressure. All have internal wire connections.

mercury-vapor-filled tubes with thermionic tungsten or oxide-coated cathodes ("sun-lamps") are used as sources of ultraviolet light.

Open arcs, which operate exposed in air, are still used for searchlights and in motion-picture projection. In the carbon arc, the emission from the cathode is thermionic in origin, arising from the intense heat present. The current flow is very great, often 100 amp./sq. cm. cross section. The ionization is corre-

spondingly intense and the bombardment of the cathode by ions results in incandescence, from which most of the light of the arc comes.

One of the latest advances is the combination of the gas discharge with fluorescence, which produces more light per watt and improves its color at the same time. Many mineral substances, notably willemite (zinc orthosilicate with silver impurities), cadmium tungstate, zinc phosphate, and zinc sulphide, emit visible light in the presence of ultraviolet light. If the inside of a gas-discharge tube is coated with a fluorescent powder, the large amount of ultraviolet light produced by the high-energy-level transitions in the gas produces a bright fluorescence in the powder, which combines with the visible light of the discharge itself. By choosing the proper fluorescent material, the color of the resultant light can be controlled over a wide range, including an almost pure white.

Problems

1. From the values of wave lengths shown in Fig. 109, compute the frequencies of the yellow light of sodium, and the limiting frequencies of the visible range of neon light and mercury light.

2. Compute the energy transition (expressed in equivalent electron volts) corresponding to the visible limits of the neon-discharge spectrum, from the wavelength data given in Fig. 109.

3. One of the lines in the sodium discharge produces yellow light of wave length 5890Å., while one of the yellow lines in the neon spectrum has a wave length of about 5700Å. The first excitation potential of sodium is 2.1 volts, of neon 16.6 volts. Assuming that the yellow light produced in the neon case results from an energy transition to the first excited level and in sodium to the normal unexcited level, compute the relative efficiency of light production of each line.

4. The light from a sodium lamp is directed toward a caesium phototube whose spectral response is given in Fig. 94. If the sensitivity of the phototube is 20 μa. per lumen for a tungsten source at 2870°K. (spectral energy given on page 180), what is its sensitivity to the light of the sodium lamp (all light output at 5890 and 5896Å.)? What is its sensitivity to neon light (assuming that all the light energy from neon is equally distributed between 5700 and 6700Å.)?

5. In a neon-sodium lamp, at operating temperature the pressure of the neon is 2.0 mm., that of the sodium 0.0002 mm. Compute the ratio of neon atoms to sodium atoms. An electron starts with an energy of 30 volts; its subsequent actions are (1) neon ionization, and (2) sodium ionization. Having ionized a neon atom what is its energy? How many elastic collisions should it make thereafter before ionizing the sodium atom, and what

energy will remain after the sodium ionization (the electron loses one part in 100,000 of its energy in each elastic collision)? How can this remaining energy contribute further to the light output of the sodium vapor?

Bibliography

See Bibliography of Chap. V (page 90).

DUSHMAN, S.: Production of Light from Discharges in Gases, *Gen. Elec. Rev.*, **37**, 260 (June, 1934).

FOUND, C. G.: Fundamental Phenomena in Sodium-vapor Lamps, *Gen. Elec. Rev.*, **37**, 269 (June, 1934).

PAULING and GOUDSMIT: "The Structure of Line Spectra," McGraw-Hill Book Company, Inc., New York, 1930.

MITCHELL and ZEMANSKY: "Resonance Radiation and Excited Atoms," The Macmillan Company, New York, 1934.

MOON, P. H.: "The Scientific Basis of Illumination Engineering," McGraw-Hill Book Company, Inc., New York, 1937.

MILLER and FINK: "Neon Signs," McGraw-Hill Book Company, Inc., New York, 1935.

CHAPTER X

SPECIALIZED ELECTRON TUBES

Introduction.—The electron tubes considered in the preceding four chapters have resulted from a straightforward application of the basic electronic principles. A source of free electrons is provided, a vacuum or gas-filled path is laid down across which they flow, and, where desired, means of controlling the flow are set up in the vacuum- or gas-filled space. Moreover, the tubes are designed to fill a wide variety of applications, *i.e.*, they are intended for general purposes within their broad fields of use. The tubes described in this chapter, in contrast, are designed with a much narrower range of applications in mind, hence they may be called "specialized" electron tubes. They are characterized by a highly ingenious, rather than straightforward, application of the basic electronic principles. They are interesting not only for this fact but because they illustrate the remarkable range of the present-day uses of the free electron and indicate some of the probable uses in the immediate future.

44. The Cathode-ray Tube.—The principle of the modern cathode-ray tube was employed, we remember, by J. J. Thomson when he determined the ratio of the electron's charge to its mass (pages 65 to 70). The principle of beam formation and deflection which he used was neglected for many years, and it was not until 1925 that the cathode-ray tube gained a place as a practical measuring device. In the short time since then, however, it has come very much into its own; it is available in a very wide variety of shapes and sizes and is applied in such a large number of ways that it hardly deserves the name "specialized" electron tube. Its basic function is, nevertheless, rather restricted. It measures the strength of a current or a voltage and is particularly valuable when the strength is varying rapidly and irregularly. Lately the cathode-ray tube has found use as a source of light for the recreation of television images; in this use its measurement ability is secondary to its ability to produce light variations with extreme rapidity.

A diagrammatic view of a typical electrostatic-deflection cathode-ray tube is shown in Fig. 118. It consists, as shown, of an "electron gun" which provides free electrons in a narrowly confined beam, and of an anode system which focuses this beam of electrons on the fluorescent screen at the other end of the tube. The deflection of the beam is obtained by applying a voltage difference to two sets of deflection plates. The envelope of the tube, usually made of glass, is in the shape of a funnel, and in the vacuum form of the tube it is very thoroughly exhausted to a pressure of about 0.001 μ (1 μ = 10^{-3} mm.). The beam-forming system consists of the thermionic cathode, usually a flat surface coated with barium oxide, which is heated by the heater wire behind it. The electrons leaving the cathode flow through

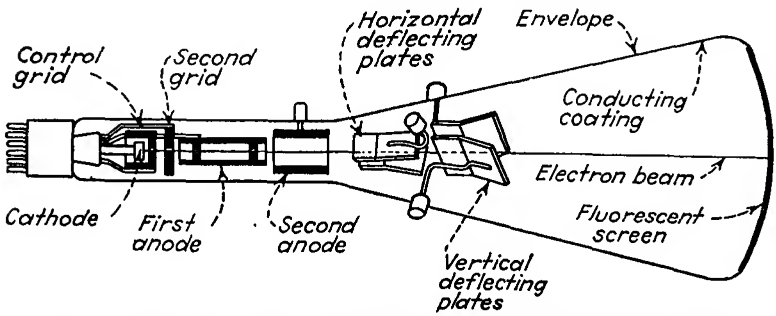


FIG. 118.—Structure of an electrostatically deflected cathode-ray tube.

an aperture, and thence to another electrode containing a second aperture. This second electrode is insulated from the cathode and serves as the "control grid" of the tube, controlling the number of electrons in the beam in accordance with the voltage applied between it and the cathode. Beyond the grid is the first anode, operated at a voltage several hundred volts more positive than the cathode, whose main function is to focus the beam by counteracting the tendency of the beam to diverge under the influence of the mutual repulsions between the electrons. Beyond the first anode is the second anode, operated several thousand volts more positive than the cathode, whose function is to focus and accelerate the electrons so that they form a narrow beam of high energy when they reach the end of the tube. At the end of their journey they impinge on the screen a white coating of willemite or some similar fluorescent

substance. The impact of the electrons on the screen has two effects: (1) It produces light of a brilliance depending on the number of electrons and their energy, and of a color depending on the type of fluorescent material used. (2) It produces a quantity of secondary electrons which must be removed from the screen. This removal is usually accomplished, in the larger size tubes, by a conducting coating of graphite inside the "funnel," which is connected electrically to the second anode.

The deflecting plates, which are parallel or slightly at an angle to each other, are provided with terminals and are insulated from the other electrodes in the tube. In the magnetic-deflection tubes, the plates are omitted and the deflection is obtained by passing current through coils whose axes are at right angles to the axis of the tube.

In some types of tube, not much used at present, inert gas is included in the tube to aid in focusing the beam of electrons. The high electron energies in the beam ionize the gas within the beam path. The resulting positive ions reduce the negative space charge, *i.e.*, the tendency of the electrons to repel one another, so that the beam maintains a narrow cross section. The gas is an unstable element in the tube, and it has other undesirable effects. In modern tubes the focusing problem is overcome by the proper design of gun and anode structures, so the gas-focusing is not necessary.

Several types of fluorescent material are in common use. The general-purpose screen is made of zinc orthosilicate (willemite); the color produced by it is green. For photographic purposes a blue color of very rapid action is desirable; cadmium tungstate is commonly used for this purpose. For certain types of visual and photographic work a long persistence of the light is valuable. This phosphorescent effect, in which the light persists for as long as several seconds after the electron beam is removed, is obtained from zinc phosphate.

Application of the Cathode-ray Tube.—The applications of the cathode-ray tube in measurement work depend on two facts: (1) when electrostatic deflection is used, a voltage may be measured without taking power from the circuit under test; and (2) it will follow changes in voltage up to 100,000,000 per second or faster. When magnetic deflection is used, current, rather than voltage, determines the deflection and hence power

is consumed. Also the upper frequency limit is determined by the inductance of the deflecting coils. Usually, therefore, electrostatic deflection is used for exacting measurement purposes.

The measurement is made in terms of the motion of the fluorescent spot across the screen. The *deflection sensitivity* (the number of millimeters deflection per volt potential difference applied between the plates) depends not only on the design of the tube but on the accelerating voltage applied to the second anode. At high accelerating voltages the sensitivity may be as low as 0.05 mm./volt. At lower accelerating voltages, the sensitivity rises as high as 0.6 mm./volt, but it rarely exceeds this amount. Since a spot motion of less than 1 mm. is hard to

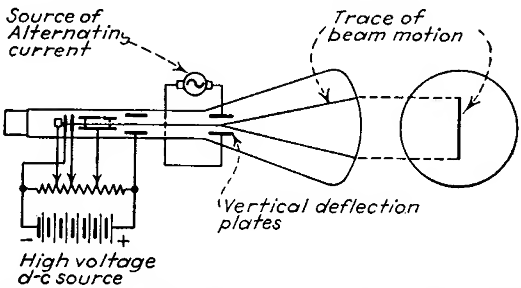


FIG. 119.—When alternating voltage is applied to one pair of the deflecting plates, the fluorescent spot traces out a straight line.

measure, this sets the lower limit of directly measurable voltage at about 2 volts. Actually much wider deflections than 1 mm. are usually desired, so that the practical voltage range lies between 10 volts and several hundred volts. If the voltage to be measured is less than this, amplification must be used.

The cathode-ray tube requires much accessory apparatus. A high-voltage power supply is essential, and a separate amplifier for each set of deflecting plates is desirable. In addition a "sweep circuit," used for the examination of alternating-current wave forms, is necessary for all but the simplest applications.

If a direct voltage is applied to one set of the deflecting plates, the resulting spot motion is a simple one-way deflection equal in amount to the sensitivity times the applied voltage. If alternating voltage is applied (see Fig. 119), the spot motion is reciprocating in form, and at all frequencies above about 10 c.p.s. the persistence of vision of the eye gives the spot the appearance of a continuous straight line. The length of this

line is a measure of the crest-to-crest amplitude of the applied alternating voltage. The line gives no indication of the *manner* in which the voltage is alternating since it does not reveal the wave form.

It is possible to show the variations of the voltage as a function of time (*i.e.*, to form a plot of voltage versus time on the screen) if the remaining set of deflection plates are called into play. For this purpose a special circuit shown in Fig. 121 is used to charge a condenser slowly and to discharge it continuously at a faster rate, the voltage appearing across the condenser

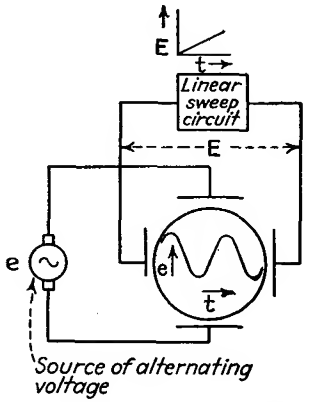


FIG. 120.—Wave-form analysis. A linear "sweep" voltage on one set of plates spreads the voltage variation applied to the other set.

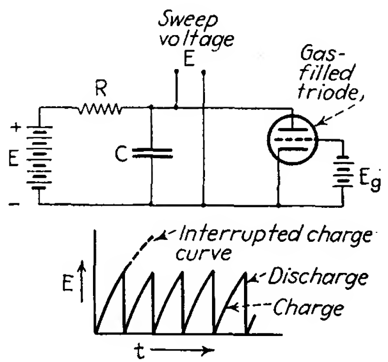


FIG. 121.—Simple sweep circuit. The gas-filled triode periodically discharges the condenser, at a rate determined by the voltage E_g.

during the charge increasing in direct proportion to the time of charge. This voltage, applied to the horizontal deflection plates, will cause the spot to move horizontally across the screen at constant speed. If the alternating voltage is applied to the vertical deflection plates at the same time, the spot will move in a vertical direction as it is moved across the screen horizontally. In this way the vertical motion is made to correspond to the measured voltage while the horizontal motion is proportional to time, and the resulting motion is a voltage vs. time plot of the measured voltage (see Fig. 120).

This type of wave-form analysis, commonly called cathode-ray oscillography, is of the greatest importance in analyzing the performance of electrical-communication equipment, in which

highly complex voltage variations are always present. The rapidity of response of the cathode-ray oscillograph is matched by no other wave-form analyzer, a fact of particular importance in radio and television research where the frequency of the current alternations is commonly as high as 50,000,000 per second.

A simple form of alternating voltage analysis, which eliminates the need of the sweep circuit, can be carried out by placing one alternating voltage on one set of plates and another on the other set. The resulting motion of the spot will be a pattern whose shape may reveal much about the applied voltages. For example, if the two voltages are of the same frequency, the same maximum strength, and are separated in phase (Chap. XI) by 90°, then the spot motion is a circle. Another form of path (Lissajous figure) is shown in Fig. 122, together with the frequency and phase relations which produce it.

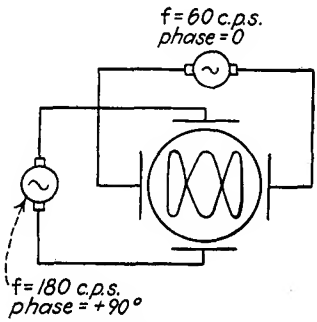


FIG. 122.—Typical Lissajous figure, produced by alternating voltages applied to both sets of deflecting plates.

The use of the cathode-ray tube in television is described in a later section.

45. X-ray Tubes.—Another member of the specialized-tube group is the X-ray tube (Fig. 123). The earliest electronic

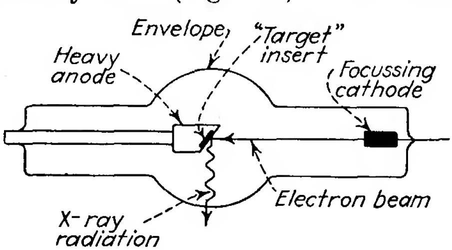


FIG. 123.—Elements of an X-ray tube.

devices to be put to practical use, X-ray tubes have been the subject of a vast amount of research and are now available for a wide variety of medical, dental, and industrial purposes. Most modern X-ray tubes are of the thermionic variety, but the earlier types contained cold electrodes and a considerable amount of gas. In the newer tubes the electron supply is taken from a thermionic

cathode, usually pure tungsten operated at about 2500°K. The anode of the tube must stand very heavy electron bombardment, since the cathode-to-anode applied voltage is from 10,000 to 500,000 in conventional tubes and up to 1,000,000 volts in advanced types.

The high applied voltage endows the free electrons with a corresponding energy. For example (since 1 electron volt is 1.6×10^{-12} erg), a 625,000-volt electron has one millionth of an erg of kinetic energy. This small energy is a huge possession for a particle weighing only a billionth of a billionth of a billionth of a gram. When such an electron hits the metal of the anode or "target," it transfers its energy to the atoms of which the metal is made. A large part of this energy is converted into heat, so in the larger tubes it is customary to employ a water-cooling system to remove the heat from the anode.

The important effect of the bombardment is the production of radiant energy, by a process similar to that employed in the production of visible light in a gas discharge, in which the atoms undergo an energy transition owing to the impact of the electrons. The energy changes involved in the X-ray tube are very great and hence the wave length of the radiation is correspondingly short. Referring to Equation (50) (page 184), it will be seen that a 100,000-volt energy change corresponds to a wave length of 0.123Å., (0.123×10^{-8} cm.). This extremely short-wave-energy length has the familiar and highly useful penetrating power which makes X-rays so valuable as a scientific and industrial tool.

A very great amount of information has been amassed concerning the use of X-rays in various pursuits, including the relation between their wave length and penetrating power, their therapeutic value, the methods of confining the radiation by lead and other heavy-metal sheaths, their use in analyzing chemical crystalline substances by refraction methods, and a host of similar items. This body of information is much too great to be examined here. The interested reader is urged to pursue the subject further in the publications listed in the bibliography at the end of the chapter.

46. The Iconoscope, the "Eye" of Television.—The iconoscope ("image viewer"), invented by Dr. V. K. Zworykin, may with justice be called the most remarkable electron tube ever invented.

It combines thermionic emission, cathode-ray-beam formation, and photosensitivity in a way which is highly ingenious, and for a very practical purpose, the perception of optical images and their conversion into an electric current. In this capacity it serves in television transmission as a camera tube for the "picking up" of the studio scene and for delivering the converted image, piecemeal, to transmission equipment.

The internal arrangement of the iconoscope is shown in Fig. 124. The lens (of the ordinary optical variety) focuses the image

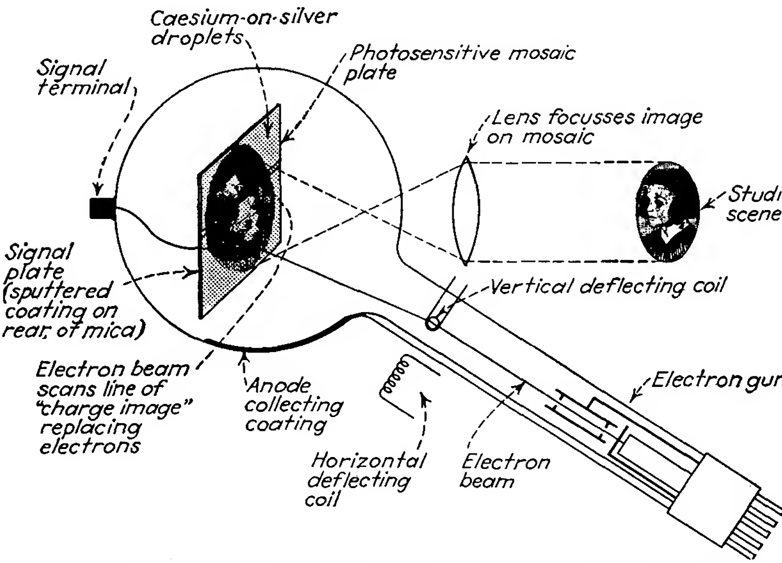


FIG. 124.—Operation of the iconoscope, a camera tube employed in television.

of the scene on the mica or "mosaic" plate. The front surface of this plate is covered with millions of tiny independent droplets of silver, each droplet insulated from its fellows by the mica. On the droplets is deposited caesium which is oxidized to caesium oxide in the same fashion as in a caesium-type phototube. Each individual droplet is thus made photosensitive, *i.e.*, capable of emitting electrons when illuminated, the number of electrons emitted being in proportion to the strength of the light and to the time of its duration. On the reverse side of the mica sheet is sputtered a thin film of metal so that the droplets and the sputtered film form a system of miniature condensers, capable

of storing charge and of developing a voltage proportional to the stored charge.

As the light from the image falls on the mosaic plate, each droplet emits electrons in proportion to the amount of illumination it receives from the image. The emitted electrons are collected by the anode coating inside the bulb and are eventually returned to the plate through the external circuit and the cathode-ray beam, as outlined below. The emission of electrons leaves the mosaic plate with a charge deficiency, and the deficiency of charge is distributed on the plate in exact reproduction of the highlights and shadows of the scene. The longer the light remains on the plate, the greater does the charge deficiency become, but because each droplet is insulated, the charge cannot rearrange itself, and hence the form of the distribution of charge over the plate is "stored."

In practice this storage process is allowed to continue for about $1/30$ sec. By this time the optical image has been converted into a sizable "charge image." The charge deficiency is then replaced by a cathode-ray beam, in an orderly sequence called "scanning." The cathode beam starts at the upper left-hand corner of the picture, moves horizontally (by means of magnetic deflection, see page 67) across the upper edge of the picture, and in so doing replaces the charge along a narrow strip. The cathode-ray beam is then extinguished by the action of the control grid of the tube and the magnetic scanning currents change in such a way that the beam is caused to move back from right to left while it is extinguished. The beam is then again released at the left-hand edge of the picture, but by this time the vertical deflecting coil system has moved the beam vertically downward a small amount, so that the beam lies just below its original position. The left-to-right "line scanning" then continues, and the beam replaces the charge in another strip lying parallel and just adjacent to the first strip. At the end of the second strip the beam is again extinguished, moved from right to left while extinguished, and then released again at the beginning of the third strip. In this manner the cathode beam covers the whole area of the mosaic plate, replacing the charge in an orderly fashion. This scanning process is in reality the conversion of the two-dimensional area of the picture into a one-dimensional series of strips, in which form they can be

converted readily into corresponding voltage or current variations.

In modern cathode-ray television practice the whole area of the picture is divided, nominally, into 441 strips, all of which are sent in $\frac{1}{30}$ sec., so that 13,230 strips are sent each second. The picture is repeated 30 times per second, so that the illusion of motion is carried, motion-picturewise, through the system. Each droplet thus has $\frac{1}{30}$ sec. in which to accumulate its charge

FIG. 125.—A form of iconoscope. The optical image passes through the flat glass window to the mosaic plate within the tube. The television signal is obtained from the two terminals at the bottom.

deficiency, whereas this deficiency is replaced in a few millionths of a second, so rapid is the action of the charge-replacing beam.

Since each droplet is one plate of a tiny condenser, the rapid replacing of the charge on it produces a rapid change of voltage between the droplet and the metallic coating behind the mica. This sudden change of voltage, appearing on the metallic coating, can be conducted to an external circuit. Furthermore, the voltage change occurring on the plate is, theoretically at least, always proportional to the charge replaced at that particular instant by the beam. Hence the voltage appearing between the

signal plate (as the sputtered coating is called) and the collecting anode changes in rapid sequence, in proportion to the successive replacement of charge on each droplet as it is scanned. The signal voltage leaving the signal plate is thus an electrical replica of the illumination changes along each line of the image. These voltage changes are amplified and transmitted by the conventional methods of radio-frequency carrier communication (Chap. XIII).

At the receiver the signal is amplified and applied to the control grid of a cathode-ray tube, on the screen of which the image is recreated. The cathode-ray beam in the receiving tube is caused to move from left to right in the same sequence as, and synchronously with, the beam in the iconoscope. At the same time the strength of the electron beam is being changed (through the radio connection) in accordance with the signal produced at the signal plate of the iconoscope. The fluorescent spot on the receiver tube thus moves across each strip of the picture, changing in brilliance as it goes, in accordance with the changes in brilliance in that portion of the original scene. The scene is thereby reproduced.

This apparently simple process is beset with the most difficult problems of any branch of the electronic art. In the first place, to eliminate flicker it is necessary to divide the scanning sequence into two groups of lines, the odd-numbered lines being sent first. The even-numbered lines then fill in the gaps, provided that they land in the proper spatial relationship to the first set. In the second place, in a 441-line picture (having a width $1\frac{1}{3}$ times its height), there are some 260,000 significant variations of light and shade. Sent 30 times a second, this means nearly 8 million significant variations to be transmitted each second, which is a specification which only the nimble electron can possibly handle. In fact, even the best electron control now available falls far short of this ideal. Time must be lost in maintaining synchronism between the beams in the iconoscope and receiver tube. The highest limit at present attained is about 4 to 6 million variations per second, which are sent by a signal whose most rapid variations are half as fast (since there are two directions, positive and negative, available for variation). Even at this comparatively low level, the results obtained under the best conditions (see Fig. 126) are remarkably good.

The Farnsworth "Image Dissector."—Another form of television camera tube used for converting an optical image into

FIG. 126.—441-line television image, as it appears on the fluorescent screen of the cathode-ray receiving tube. (*Photo taken in the experimental television studios of the National Broadcasting Company.*)

an electric signal is the Farnsworth image dissector. In it a uniform photosensitive plate is used, rather than a mosaic. At

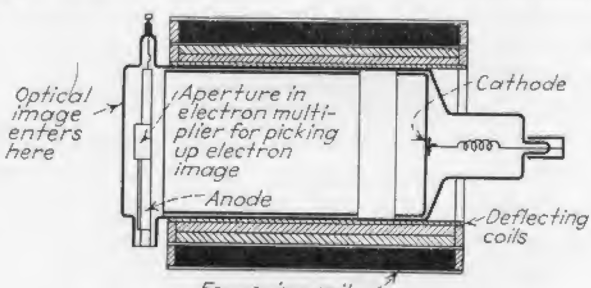


FIG. 127.—Farnsworth image-dissector tube.

one end of the tube (see Fig. 127) is a flat silver cathode covered with a caesium photosensitive layer, on which the optical image

is focused by means of a lens. The whole plate thus emits electrons, the number of electrons freed from a given point being proportional to the illumination at that point. All of the electrons are attracted, en masse, to an anode coating at the opposite end of the tube. As this "picture in electrons" travels down the length of the tube, it tends to diverge, under the repulsion forces acting between the electrons. This tendency is curbed by the application of an external magnetic field which restricts any radial motion and brings the electron picture to focus at the anode end of the tube. By means of still other magnetic deflecting fields, the picture in electrons is caused to move bodily, horizontally and vertically, past a small aperture leading to an auxiliary chamber in the tube. This aperture thus selects electrons from the electron image in the proper scanning sequence, and directs them to a current-intensifiying device, or "multipactor," a form of electron multiplier which depends on secondary emission. The electron multiplier is used to increase the strength of the electron current, an increase which is needed since the tube does not possess the storage properties of the iconoscope. From the electron multiplier the electron flow is caused to pass through a resistance, across which the signal voltage appears and from which it is conducted to the transmitting equipment.

47. Electron-image Tubes.—The same basic idea of forming an electron image has been used in the electron telescope, a form of tube developed by Dr. V. K. Zworykin and his associates. This tube, developed in connection with research on the iconoscope, promises to have many uses far removed from television. In particular it is able to transform images in invisible infrared light directly and instantaneously into visible images, and to magnify or reduce their size at the same time.

The tube (Fig. 128) is cylindrical in shape, and has flat surfaces. Inside one face is a transparent glass surface on which is sputtered a caesium-oxide-on-silver photosensitive cathode surface. At the other end of the tube is a fluorescent screen of willemite, coated on the inside of the flat surface of the tube. Between the cathode and the screen are a series of ringlike anodes and an aperture, which direct electrons leaving the cathode to the screen. In use, light (visible or infrared) is allowed to fall on the cathode surface, in the form of an image focused on it through

a lens system. The electrons leave the cathode as an electron image and are focused (by the electrode system) on the screen, which reproduces the electron image in visible light. By changing the voltages applied to the ring and aperture electrodes it is

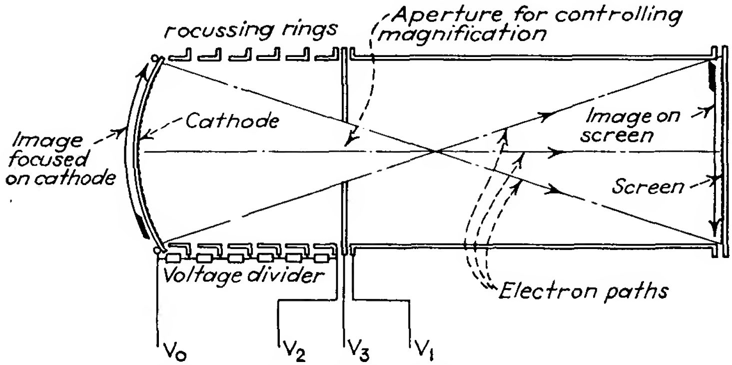


FIG. 128.—Element structure of the "electron telescope," an electron-image tube.

possible to change the paths taken by the electrons, so that they diverge or converge. The divergence results in a magnification of the image (hence the name "telescope") which can be controlled electrically; the convergence produces a reduction in image size. The electrode structure (rings and aperture) thus acts on the electron paths in the same way that a lens acts on rays of light; the electrode structure is, in fact, commonly called an electron lens and is designed by careful study of the principles of electron optics, that is, the study of the influence of electrostatic and magnetic fields on the paths pursued by fast-moving electron beams.

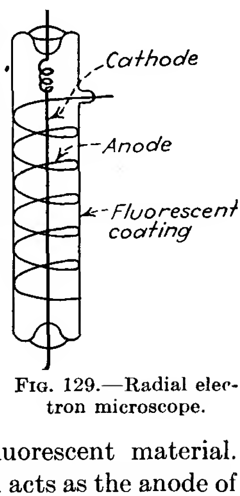


FIG. 129.—Radial electron microscope.

A Simple Form of Electron Microscope.—A most simple application of electron optics is made in the electron microscope shown in Fig. 129. The cylindrical tube is coated on the inside with willemite or some other fluorescent material. Against this coating rests a coil of wire which acts as the anode of the tube. Down the center of the tube, on the axis of the cylindrical envelope, runs a straight tungsten wire which acts as a

thermionic cathode, operated at 2500°K. by passing a heating current through the wire. When a high voltage is applied between cathode and anode, the emitted electrons travel radially from the cathode wire to the fluorescent anode coating and there produce a visible image. This image is found to be a more or less exact reproduction of the appearance of the surface of the wire viewed under an optical microscope. The tube is thus a form of electron microscope, its magnification being equal to the ratio of the diameter of the cylindrical envelope to the diameter of the wire, (several hundred times in practical tubes). The explanation of the action of the tube is given by the fact that the fluorescent image is actually a reproduction of the variations of electron emission over the surface of the cathode, which in turn are produced to a large degree by local variations in the surface of the wire. Die marks made in drawing the wire, for example, produce local sharp ridges from which the emission is higher than average (doubtless, part of this increase in emission is due to field emission, since the strength of the electric field at each sharp point on the surface is much greater than average). These die marks show up in the magnified image. If they are removed by careful polishing of the wire, their image disappears from the screen.

The same technique has been used, in much more elaborate electron-microscope tubes, for examining the crystal boundaries in metallic cathodes, and even for magnifying the structure of very thin biologic specimens through which a diverging beam of electrons is caused to pass. The whole subject of electron optics is new, and its mathematical development is very complex. But it can be expected to serve as an important research tool as the technique of applying it is improved. Incidentally, it is worth noting that the design of cathode-ray tubes, both the conventional varieties and the specialized photosensitive types used in television, is essentially an electron-optical problem, solved in much the same way as a problem in the optics of light.

48. Electron-multiplier Tubes. Secondary Emission Applied. Whenever electrons of 30 volts or higher energy strike a metallic surface of low work function, secondary electrons are emitted, and usually more than one secondary electron is released by every impacting electron. This effect is a nuisance in conventional tetrode tubes, as we have seen, and has led to the

inclusion of the suppressor grid between the screen grid and the plate. But secondary emission has been put to use in several specialized tubes, called electron multipliers. One use of these multipliers has already been mentioned, in connection with the Farnsworth image dissector tube.

Another form of multiplier tube developed by Farnsworth is shown in Fig. 130. At either end of the tube are caesium-coated plates which will liberate from 2 to 10 secondary electrons for every impinging electron. Across the plates is applied a source of alternating voltage of very high frequency, such that the time of one-half cycle is approximately equal to the transit time of the electrons traveling between the end plates. Midway

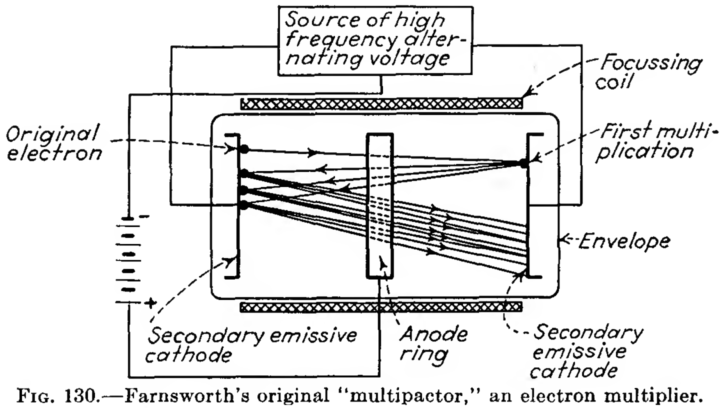


FIG. 130.—Farnsworth's original "multipactor," an electron multiplier.

between the plates is a ringlike anode. An original electron, perhaps liberated from the caesium by the action of light, is attracted toward the anode but is prevented from hitting it by the action of an external magnetic field, applied by the coil shown. The electron travels, therefore, to the opposite plate which it hits with sufficient energy to liberate several secondary electrons. By this time the direction of the applied alternating voltage has reversed, so the newly freed electrons are repelled toward the center anode, which cannot collect them. The new electrons thus travel to the opposite end plate and there liberate still more electrons. The number of electrons participating in the flow thus builds up cumulatively until limited by space charge, at which point an equilibrium level is reached, the anode collecting a steady current composed of the slower

velocity electrons which elude the regulatory action of the magnetic field. Once this tube is started, it will supply its own alternating voltage, the power being taken from the battery connected in series with the anode; hence it can be used as a cold-cathode supplier of alternating-current energy. Other forms of Farnsworth "multipactor" tubes are built with a single cylindrical cathode, coated with caesium on the inner surface, across the diameter of which the electrons fly. Such tubes have been demonstrated producing many hundreds of watts of alternating power at high efficiency.

A somewhat different form of electron-multiplier tube has been developed independently by Zworykin and Farnsworth. The form due to Zworykin is shown in Fig. 131. Inside a cylindrical

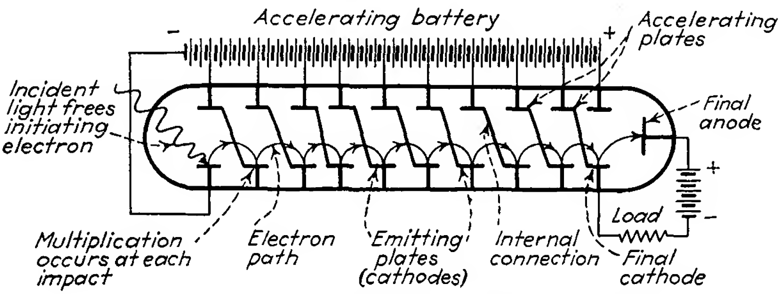


FIG. 131.—Dr. Zworykin's multistage electron multiplier.

tube are two sets of plates, the lower set treated with caesium to emit secondary electrons, the upper being merely accelerating anodes. A source of high voltage is connected across the first cathode and the last anode as shown, the intermediate anodes and cathodes being connected in series. Each anode is thus at a positive potential with respect to the cathode immediately opposite it. Light shines on the first cathode, freeing an electron which is attracted to the opposite anode. An external magnetic field forces the electron to travel in a circular path, however, so the anode does not collect the electron. Instead it hits the second cathode, there liberating several secondaries. These in turn travel in a circular path to the next cathode, liberating still more secondaries, and so on. This process continues from cathode to cathode until finally the last cathode is reached, and the resulting electron stream is then collected by the final anode and thence conducted to the external load. If the tube contains

10 cathodes, and if each impinging electron releases 5 secondaries on the average, then the original electron is multiplied

$$5^9 = 1{,}953{,}125 \text{ times.}$$

The electron current entering the last anode is then nearly two million times as strong as that leaving the first cathode. This is an enormous current amplification for so compact a tube. Its uses in the future of electronics may be of considerable importance. If, for example, light from the sound track of a motion-picture film is allowed to fall on the first cathode, the current flow throughout the tube will follow the changes in illumination produced by the variations in the sound track, and the output current will be strong enough to operate a loud-speaker. One such tube has, in fact, filled an auditorium capable of seating 2000 people. Compare this simple tube with the roomful of conventional apparatus now used in theaters to convert the sound track into sound.

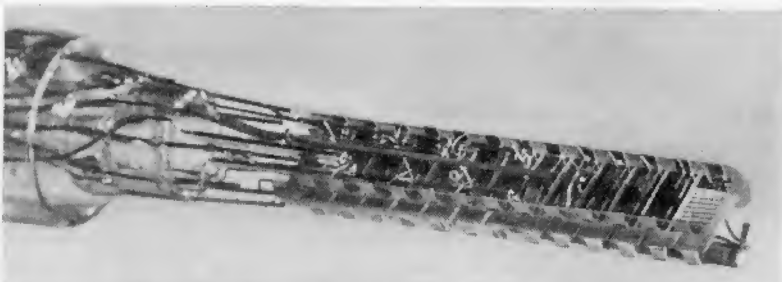


FIG. 132.—Element structure of a 12-stage multiplier.

The Dynatron.—A characteristic of the tetrode tube (and of the triode tube operated with a positive grid) is the collection of a secondary emission current by the positive grid. Under certain conditions, this secondary emission current will be larger than the primary plate current, and the total plate current will *decrease* as the plate voltage is increased. This effect is shown in Fig. 70, page 125, at plate voltage between 20 and 60 volts. The cathode-to-plate path in the tube has, in other words, a negative dynamic plate resistance, which may be used to neutralize the ordinary positive resistance of the external circuit. In such a circuit, whose net resistance is zero, oscillatory (alter-

nating) currents will flow without diminution, and hence the tube may be used as a source of alternating current and voltage. So used, the tube is called a *dynatron*, as named by its inventor, A. W. Hull.

The Movable-anode Tube.—The current flow in a vacuum diode tube depends not only on the voltage applied between the electrodes but also upon their physical separation (the Langmuir-Child law, page 59). A new form of tube, shown in Fig. 133, is contained in a metal envelope, one end of which is thin enough to be flexible. Through this thin section is passed a rod, on the inner end of which the anode of the tube is mounted. If the external projection of the rod be moved, the separation of the anode and cathode can be varied, and the current flow (which varies inversely as the square of the separation) is thus changed over a wide range. A sensitive milliammeter connected in the cathode-anode circuit will register the changes in current, and thus can be calibrated directly in distance units. The tube has been used as a direct-reading electric micrometer and in other mechanical applications in which small displacements must be measured.

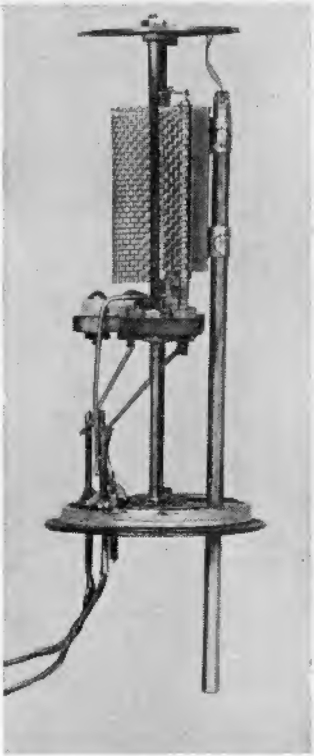


FIG. 133.—Elements of the movable-anode tube.

49. Specialized Gas-filled Tubes. *The Ionization Gauge.*—The thermionic ionization gauge is used to measure very low pressures, as low as 10^{-9} mm., and hence hardly deserves to be called a "gas-filled" tube. It operates on the principle of ionization, however, and depends therefore on those few molecules of gas which are present at low pressures. The tube is essentially a triode, but the plate is operated at a negative potential with respect to cathode, and the grid at a positive potential, as shown in Fig. 134. Electrons leaving the cathode (usually a pure-tungsten or thoriated-tungsten wire, since it

liberates but little gas during life) are attracted by the near-by positive grid, and in so doing encounter neutral molecules of gas, ionizing them. The positive ions so produced are attracted by the negative plate, are collected by it, and so produce an ion current which can be measured by a meter in the plate circuit. The strength of this ion current is in direct proportion to the number of ions reaching the plate, which in turn is proportional to the number of neutral molecules ionized. The ion current is then a measure of the number of molecules present and hence of the gas pressure within the tube. Such ionization gauges are commonly used in the pumping of large tubes for transmitting and X-ray use, as a continuous indicator of the vacuum conditions attained.

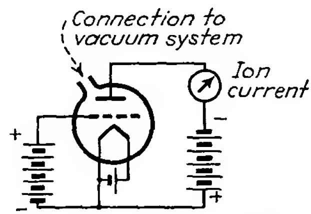


FIG. 134.—Ionization gauge.

The Strobotron, a Stroboscopic-light Source with Control Abilities. Stroboscopic light is light which is regularly interrupted at a rate corresponding to the rate of rotation, reciprocation, or oscillation of an object to be examined by the light. When so examined, the object appears to the eye to be motionless, and can be examined at length, many characteristics of the motion being visible which would otherwise completely escape detection. Electronic methods are employed to secure a controllable rate of interruption. The strobotron, a stroboscopic-light source developed by Germeshausen and Edgerton, is a neon-filled tube containing a caesium-coated cold cathode on which forms a cathode spot similar to that of the mercury-pool type of cathode. The tube (Fig. 135) contains two grid structures, between which the control voltage is applied to start the discharge. Once started, the discharge transfers to the cathode and anode, producing a brilliant light, and a current flow which can be used for control purposes. Originally developed for use in a semiportable stroboscope, the tube has found use in

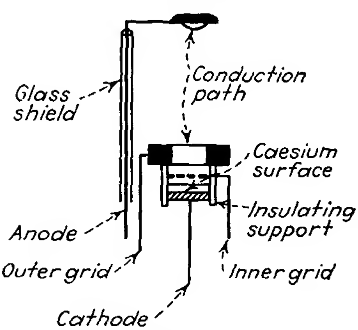


FIG. 135.—Elements of the strobotron.

welding control, and similar applications where the light produced is of no consequence.

Since no heating current is required for the cathode, the over-all efficiency of the tube is high. It can readily be operated from batteries, a great convenience in portable apparatus and in isolated locations. The two grids can be used to exert a dual control, such that the discharge starts when the voltages applied between them and the cathode reach a definite difference, *i.e.*, control may be obtained in terms of two voltages instead of one. The characteristics of the tube's action have been examined at length, although its control applications are still in the development stage.

Problems

1. A 60-cycle alternating-current wave is applied to the horizontal deflecting plates of a cathode-ray tube. A 120-cycle wave of the same amplitude is applied to the vertical deflection plates. At the start, the maxima of the two waves coincide. By graphical means construct the resulting path (Lissajous figure) of the fluorescent spot.

2. A sweep circuit voltage is applied to the horizontal plates of a C-R tube. The sweep time is $\frac{1}{100}$ sec., and the return time is $\frac{1}{1000}$ sec. A 60-cycle wave is applied to the vertical plates, the maximum occurring at the instant the sweep begins. By graphical construction, draw the resultant motion of the fluorescent spot during the succeeding 0.033 sec. (three complete sweeps and returns). What is the necessary condition if successive cycles of the alternating-current wave are to occupy the same position on the screen?

3. To penetrate a certain thickness of lead, X-rays of 0.04Å. wave length are required. Calculate the voltage required across the cathode and anode of the X-ray tube. Calculate the velocity of the electrons as they hit the anode, (*a*) neglecting the Einstein-Lorenz law, and (*b*) taking the correction into account, assuming that the electrons start at the cathode from rest and neglecting space charge. If the anode current is 0.5 ma. and the efficiency of X-ray conversion is 1 per cent, what amount of heat must be dissipated by the anode in 1 sec.?

4. Television pictures have gone through the following stages of development: 16 lines per picture, 24 lines, 60 lines, 120 lines, 240 lines, 343 lines, 441 lines. The pictures were sent 24 per second in all cases except 343 and 441 lines, which are sent at 30 per second. The picture width is $1\frac{1}{3}$ times its height. Assuming that the detail along each line is equal to the detail at right angles to each line, derive a formula for the number of essential picture elements to be sent each second, and calculate the number in each of the seven cases given above.

5. In an electron microscope similar to that shown in Fig. 129, it is found that during 1000 hr. of operation a certain horizontal linear dimension in

the image has increased in size by 15 per cent. Find the corresponding filament-wire size (after evaporation), if the tube is 2 in. in diameter and the wire when new was 0.01 in. in diameter. Calculate the weight of the evaporated tungsten (specific gravity 18.8 g./cc.) if the effective length of the cathode is 10 in.

6. Find the negative dynamic resistance (between 20 and 60 volts) of the tetrode whose characteristics are given in Fig. 70, when $e_{c1} = -1.5$ volts.

7. The separation between cathode and anode in a movable anode tube ranges from 0.05 to 0.5 cm. At minimum separation the current flow, with 100 volts applied, is 15 ma. Draw a plot of current against lever displacement, if the effective lever ratio is 5 to 1.

Bibliography

BATCHER, R. R.: Applications of the Cathode Ray Oscillograph, *Proc. I.R.E.*, **20**, 1878 (December, 1932).

"Cathode-Ray Tubes and Allied Types," Technical Series TS-2, RCA Manufacturing Company, 1935.

RIDER, H. H.: "The Cathode Ray Tube at Work," Rider Publishing Company, 1936.

SMITH, H. X.: Cathode Ray Oscillographs, *Bell Sys. Tech. Jour.*, **20**, *YYY* 1933.

CLARK, G. L.: X-rays—What Should We Know About Them? *Elec. Eng.*, No. 85, (January 1935).

CLARK, G. L.: "Applied X-rays," McGraw-Hill Book Company, Inc., New York, 1932.

TERRILL and ULREY: "X-ray Technology," D. Van Nostrand Company, Inc., New York, 1930.

ST. JOHN and ISENBURGER: "Industrial Radiology," John Wiley & Sons, Inc., New York, 1934.

GLASSER, O.: "The Science of Radiology," C. C. Thomas, 1933.

ZWORYKIN, MORTON, and FLORY: Theory and Performance of the Iconoscope, *Proc. I.R.E.*, **25**, 1071 (August, 1937).

ZWORYKIN, V. K.: The Iconoscope—A Modern Version of the Electric Eye, *Proc. I.R.E.*, **22**, 16 (January, 1934).

"Television," vols. I and II (Collected RCA Papers) RCA Institutes Tech. Press (1936, 1937).

The Electron Telescope, *Electronics*, **9**, 10, (January, 1936).

Electronic Self-portraits, *Electronics*, **10**, 22, (March, 1937).

An Electron Multiplier, *Electronics*, **7**, 242, (August, 1934).

Secondary Emission Electron Multipliers, *Electronics*, **8**, 10 (November, 1935).

GAGER and RUSSELL: Quantitative Study of the Dynatron, *Proc. I.R.E.*, **23**, 1536 (December, 1935).

MACARTHUR, E. D.: Movable Anode Tubes, *Electronics*, **10**, 16 (March, 1937).

GERMESHAUSEN and EDGERTON: The Strobotron, *Electronics*, **10**, 12 (February, 1937); **10**, 18 (March, 1937).

PART III

ELECTRON-TUBE APPLICATIONS

CHAPTER XI

ELEMENTS OF CIRCUIT THEORY AS APPLIED TO ELECTRON TUBES

Introduction.—Electron tubes are useless unless they are connected with other pieces of electrical apparatus, that is, unless they are part of an electric circuit. The study of electronic applications is, therefore, the study of the types of electric circuits in which electron tubes are used. The principles involved in these circuits are, of course, those underlying all electrical engineering, but the ways in which they are applied reflect the extraordinary capabilities of the electron tube. This fact places emphasis on aspects of circuit analysis, particularly the study of the tuned circuit and of complex alternating currents, which are not ordinarily stressed in the conventional electrical-engineering course. The present chapter has been written to emphasize principles of value in understanding electronic-circuit functions. For the reader who desires a broader view, the references listed at the end of the chapter will be of value.

50. The Current-voltage Relationships in Electric Circuits.—An electric circuit serves primarily to direct energy from one piece of apparatus to another. In power practice the circuit receives energy from a generator and transfers it to a device, such as a lamp or motor, which consumes the energy. In communications circuits, in which electron tubes find their widest use, the transfer is of a more subtle type; it is the transfer of electrical energy in a form suitable for conveying intelligence. In either case, the utility of the circuit for performing a given task is analyzed by following the energy from its source to the final receiver and by noting the increases or decreases in energy level which occur en route.

The energy flow is examined in terms of the current-voltage relationships which exist in each piece of apparatus or "element" connected in the circuit. If the voltage across a circuit element and the current flow through it are known, then the energy

flow to or from this element can be calculated by forming the product of the voltage, the current, and the time during which the current flows. The time factor is usually common to all the circuit elements and hence may be "canceled out," leaving current and voltage as the determining factors.

In practical terms the problem may be stated in two ways: (1) Given a circuit element, to find the current flow through it when a certain type of voltage is applied; or (2) given a circuit element, to find the voltage which appears across it when a given type of current flows through it. In either case, the current and the voltage are always associated with each other, hence it is convenient to take their ratio and to use this ratio as a term descriptive of the circuit element in question. When an applied voltage e causes a current flow i in a particular circuit element, then the ratio e/i is called the *impedance* of that circuit element and is given the symbol Z.

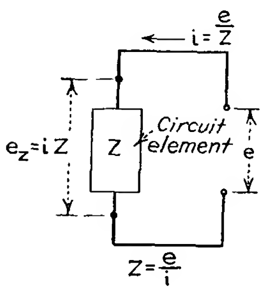


FIG. 136.—Relationships of current, voltage, and impedance, in a two-terminal circuit element.

The concept of impedance (which includes the more familiar terms "resistance," "inductance," and "capacitance") is highly useful. In the first place, it permits dealing with a circuit element, or any combination of them, without reference to any particular applied voltage or current. A circuit may be designed and its performance predicted thereby, in general terms, without reference to any particular operating condition and hence applicable to all possible operating conditions. In the second place, the effects of connections between circuit elements can be computed in terms of their combined impedance values much more simply than if the individual voltages and currents are considered.

The impedance Z of a circuit element (see Fig. 136), then, is the ratio of the applied voltage e to the resulting current i, and is given in the generalized Ohm's law

$$Z = \frac{e}{i}. \tag{51}$$

Z is in ohms when e is in volts and i in amperes. This definition-equation may be used, whenever two of the quantities are known,

to find the third in terms of the other two. But the value of Z depends not only on the type, size, and form of circuit element, but also on the type of applied current or voltage. It would appear, therefore, that we must consider all the possible types of circuit element (to name but a few: resistors, coils, condensers, transformers, electron tubes, batteries, meters, generators, motors, relays, lamps, loud-speakers, and microphones, each of which is available in a large number of different types), and that we must apply to these units all conceivable types of applied voltage. Fortunately this infinity of experiments need not be performed. Workers in electric-circuit analysis have succeeded in reducing the number of essential circuit elements to three and the number of essential types of applied voltage to one. The three essential circuit elements are *resistance* (R), *inductance* (L), and *capacitance* (C). The single essential applied voltage is the sinusoidally-varying alternating voltage. Every known type of practical circuit element may be reduced to a combination of R, L, and C, and every known type of applied voltage may be reduced to a combination of sine-wave alternating-current components, at least with sufficient accuracy for engineering purposes. Sometimes the combinations of L, C, and R and of sine-wave alternating-current components are very complicated, so that short cuts are necessary. But the basic scheme involves only three items: (1) setting up the basic alternating voltage; (2) evaluating the current produced by this voltage in the basic elements R, L, and C; and (3) setting up rules for combining the R, L, and C into circuit elements and combinations actually used in practice. If the actual applied voltage is a combination of several sine-wave components, then each is treated separately. In this way a problem in circuit analysis which at first glance may seem hopeless is quickly reduced to a matter of algebraic manipulation. In later paragraphs several examples of this process are given.

51. The Basic Sine-wave Voltage.—The basic sine-wave voltage lends itself best to graphical description. In Fig. 137 are shown two cycles of such a voltage, the instantaneous amplitude of the voltage appearing across two terminals plotted against time. This plot shows that the voltage between the terminals changes polarity regularly, and that between changes the voltage goes through zero (disappears altogether). It

will be noted also that the change in voltage is most rapid when it goes through zero and least rapid when it has its maximum or minimum values.

It may well be wondered why such a complicated type of varying voltage is selected as basic in circuit analysis. There are several very good reasons. In the first place, this type of voltage variation (or at least a very close approximation to it) occurs commonly in power practice, as, for example, in 60-cycle distribution circuits. A much more potent reason was discovered by Fourier, the French mathematician, who showed

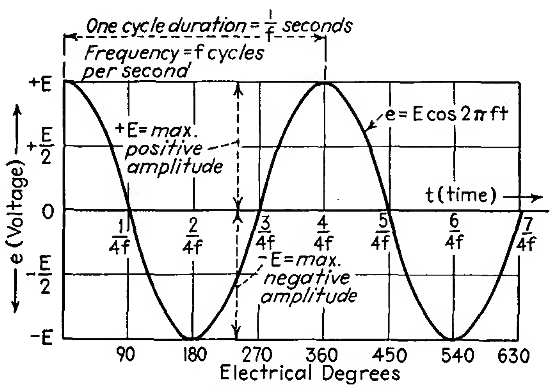


FIG. 137.—Sinusoidal alternating voltage.

that *any* type of voltage variation, no matter how bizarre, can be analyzed into a sum of several components, each of which has the form of the sine wave. Thus a knowledge of the sine wave and its effects, plus a knowledge of Fourier analysis suffice to deal with any applied voltage. In the third place, the sine-wave alternating voltage is one of two forms[1] of voltage variation which will produce the same form of current variation in all three basic circuit elements, R, L, and C. This latter fact makes it possible to calculate the current flow in any combination of circuit elements, caused by any applied voltage, simply by treating each sine-wave component of the applied

[1] The other type of voltage variation which has this property, the exponential transient, is not widely useful because it is not continuous, *i.e.*, the variation of voltage eventually comes to an end, often in a very short time. This transient effect is of interest in connection with sudden pulses of voltage in certain types of thyratron circuits referred to on page 265.

voltage separately and adding together all the resulting sine-wave currents.

The Cosine Formula for the Sine-wave Voltage.—The form of the voltage variation shown in Fig. 137 may be plotted directly from the following equation:

$$e = E \cos 2\pi ft, \tag{52}$$

where e is the instantaneous value of the voltage at time t, f is the *frequency, i.e.*, the number of complete cycles undergone by the voltage in a second, and E is the maximum or minimum *amplitude* of the voltage variation. Since the sine-wave form is

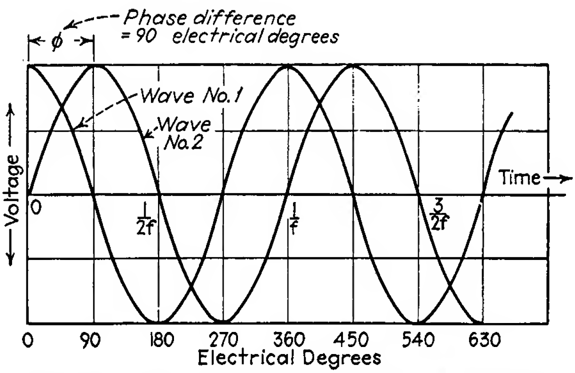


Fig. 138.—Phase relationship between two alternating voltages of the same frequency.

always assumed, only two factors, the frequency and the amplitude, serve to specify any single voltage. The amplitude E states how great the voltage becomes and the frequency f states how often it gets that great. However, if two sine waves (two voltages, two currents, or a voltage and a current) are present simultaneously, then one additional factor is required to specify the relationship between the two. This factor is the *phase*, the time separating the maximum of one wave from an adjacent maximum of the other (see Fig. 138); it is measured usually in electrical degrees, 360 electrical degrees being equal to the time of one cycle, which in turn is numerically equal to the reciprocal of the frequency

$$360° = \frac{1}{f} \text{ sec.} \tag{53}$$

These three items, *amplitude*, *frequency*, and *phase*, serve to specify any sine-wave voltage, or sine-wave current resulting from such a voltage, or any combination of such currents or voltages. Since we have agreed to restrict ourselves to the sine-wave form, we can henceforth dispose of Fig. 137, except for illustrative purposes, and use the numerical factors, amplitude (E volts or I amp.), frequency (f cycles per second), and phase (ϕ electrical degrees).

The reader may wonder where the familiar and useful direct current fits into the picture just presented. A direct current does not vary. Hence we may permit ourselves to state, as substantially equivalent, that a direct current varies no times per second, *i.e.*, its frequency is zero. If $f = 0$ is substituted in Equation (52), the result is (since $\cos 0 = 1$)

$$e = E \tag{52a}$$

In other words, the instantaneous value of the voltage e does not change but has a constant value E, the steady amplitude of the direct voltage. This device of calling a direct voltage an alternating voltage of zero frequency is convenient also in connection with the impedance formulas given in the next section.

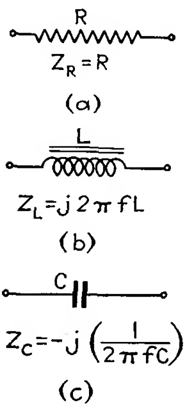


FIG. 139.—Symbols and impedance formulas for resistance, inductance, and capacitance parameters.

52. The Current-voltage Relationships of R, L, and C.—It now remains to state the current flow which results when the sine-wave alternating voltage is applied to the basic circuit elements, resistance, inductance, and capacitance. The current flow will have a sine-wave form, and it will have the same frequency as the applied voltage, hence the only questions to be answered are the amplitude of the current and its phase with respect to the applied voltage. The amplitude and phase depend on the type of circuit element in question (that is, R, C, or L) and on the frequency of the applied voltage.

The Resistance Parameter.—Electrical resistance (Fig. 139*a*) is the type of circuit parameter which converts part of the energy flow through it into heat, although it is more commonly used in electronic circuits simply to convey current into different paths or conversely to develop a voltage drop from a current flow.

If the basic voltage $e = E \cos (2\pi ft)$ is applied to a resistance R ohms, the current flow i through the resistance will be

$$i = \frac{E}{R} \cos (2\pi ft). \tag{54}$$

The amplitude I of the current flow is thus E/R amp. when E is in volts and R is in ohms. The phase difference between the voltage and the current is zero, that is, the maximum current and the maximum voltage occur at the same time.

The relationship between the amplitude of the applied voltage E, the amplitude of the resulting current I, and the resistance value R is

$$I = \frac{E}{R}. \tag{55}$$

Usually it is assumed that the resistance value does not depend on frequency. In the high frequencies commonly encountered in electronic circuits, the value of R does depend on frequency, however, so that it is necessary to specify the resistance value at the operating frequency.

The impedance Z of a resistance is by definition e/i. Hence, by Equations (52) and (54),

$$Z_R = \frac{e}{i} = R \tag{56}$$

The impedance of a resistance is thus its resistance value in ohms, specified at the operating frequency.

The Inductance Parameter.—Electrical inductance (Fig. 139*b*) is the type of circuit parameter which stores energy in a magnetic field, and hence commonly takes the form of a coil of wire wound on a magnetic core. Inductance is used in frequency-responsive circuits and for current-limitation purposes. If the sine-wave voltage of Equation (52), $e = E \cos (2\pi ft)$, is applied to an inductance of L henrys, the current flow i will be:

$$i = \frac{E}{2\pi fL} \cos (2\pi ft - 90°). \tag{57}$$

The amplitude of the current I is then:

$$I = \frac{E}{2\pi fL}. \tag{58}$$

The phase difference between the maximum of voltage and the maximum of current is 90°, (one-quarter cycle) and the maximum of current *lags behind* the voltage. The impedance of an inductance is, from Equations (52) and (57),

$$Z = \frac{e}{i} = 2\pi fL. \tag{59}$$

The impedance thus increases as the frequency of the applied voltage increases, a fact which makes inductance of value in frequency-responsive circuits.

To express the fact that the phase difference between voltage and current in an inductance is 90°, the ratio of voltage to current is commonly called inductive reactance and is given the symbol jX_L. The symbol j ($j = \sqrt{-1}$) indicates a 90° phase difference and is carried through the algebraic manipulation as outlined below. Hence, Equation (59) for the impedance, including the phase displacement, is written:

$$Z_L = \frac{e}{i} = jX_L = j2\pi fL. \tag{60}$$

The Capacitance Parameter.—Electrical capacitance (Fig. 139*c*) is the type of circuit parameter which stores energy in an electric field, the field being commonly formed between a system of conducting plates separated by an insulator, and called a *capacitor*. When an alternating voltage is applied between two such insulated plates, an apparent current flow takes place through the insulation, owing to the fact that the electrons in the current flow are alternately stored on each conducting plate, on successive halves of each alternating-current cycle. Ordinarily there is no passage of electrons through the insulation, but the passage of electrons through the external circuit connected to the two plates, as the stored charge passes from one plate to the other, is a true alternating current. The amount of current depends on the frequency, hence condensers are used in frequency-responsive circuits.

When the sine-wave voltage [Equation (52)] is applied to a condenser of C farads, the resulting current flow is

$$i = E2\pi fC \cos (2\pi ft + 90°) \tag{61}$$

The amplitude of the current is

$$I = E2\pi fC \tag{62}$$

and the phase difference between voltage and current is 90°, but in contrast to the inductance, the maximum of current precedes the corresponding maximum of voltage, *i.e.*, the current leads the voltage by one-quarter cycle. To express this fact the ratio of voltage to current is called capacitive reactance and is given the symbol $-jX_C$. The symbol j is used as in the inductive case, except that the opposite sign is used, to express the

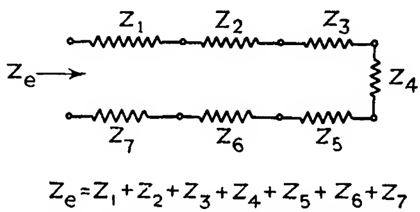


FIG. 140.—Impedances in series.

reversal of phase. Hence the impedance of a capacitance is, in correspondence to Equation (60)

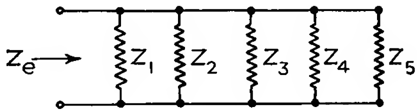
$$Z_C = \frac{e}{i} = -jX_C = \frac{-j}{2\pi fC}. \tag{63}$$

The impedance of a capacitance thus decreases as the frequency increases, and furthermore it decreases as the value of the capacitance C increases. For direct current ($f = 0$) the impedance is infinite, that is, no direct current can flow through a condenser.

By means of Equations (56), (60), and (63) it is possible to compute the impedance of a resistance, inductance, or capacitance in terms of the frequency of the applied voltage and the value of the parameter itself. This impedance value gives the amplitude of the resulting current, and, in terms of the j symbolism, the phase of the resulting current with respect to the applied voltage. When only isolated parameters are present, these three equations suffice to calculate circuit performance. More commonly, however, the parameters R, L, and C are found in combination, and it is then convenient to evaluate their *combined effect* by combining the impedances of each into one equivalent

impedance which will give the amplitude and phase of the current flow in the combination.

53. The Rules for Combining Circuit Impedances.—The circuit parameters L, R, and C are shown combined in a complex



$$\frac{1}{Z_e} = \frac{1}{Z_1} + \frac{1}{Z_2} + \frac{1}{Z_3} + \frac{1}{Z_4} + \frac{1}{Z_5}$$

FIG. 141.—Impedances in parallel.

circuit in Fig. 142a. The dotted lines represent divisions of this circuit into parts which are to be dealt with separately. The parts enclosed by dotted lines are similar to the two basic connections shown in Figs. 140 and 141, the series circuit and the

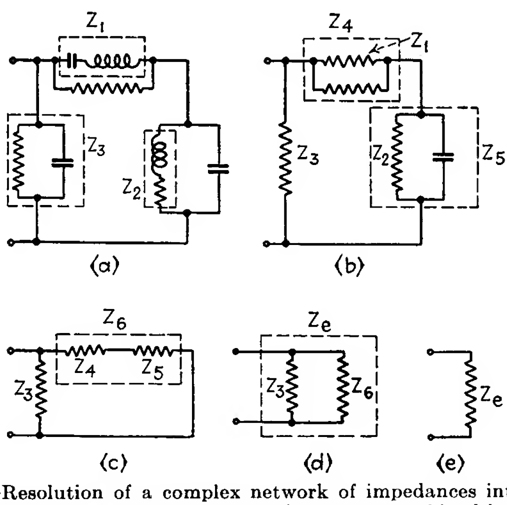


FIG. 142.—Resolution of a complex network of impedances into an equivalent impedance. The groups within dotted lines are combined by the methods illustrated in Figs. 140 and 141.

parallel circuit. Each of these basic connections can be reduced to a single equivalent impedance Z_e by the following rules:

From the series circuit:

$$Z_e = Z_1 + Z_2 + Z_3 + Z_4, \text{ etc.} \tag{64}$$

For the parallel circuit, a reciprocal relationship is necessary:

$$\frac{1}{Z_e} = \frac{1}{Z_1} + \frac{1}{Z_2} + \frac{1}{Z_3} + \frac{1}{Z_4}, \text{ etc.} \tag{65}$$

When each group has been thus reduced to an equivalent impedance, the equivalent impedances themselves may be grouped into an impedance equivalent to that of the several groups together. In this way, step by step, the entire complex maze of impedances is reduced to a single value. This single value then serves to indicate what current will flow between the terminals in question when a given voltage is applied, or conversely, what voltage will appear between the terminals when a given current flows through the impedance group.

Fortunately in electronic circuits it is seldom necessary to consider a circuit as complex as that of Fig. 142. In particular, it is often possible to consider the grid and plate circuits of a triode (or the equivalent triode of a tetrode or pentode) as isolated circuits, each of which is usually made up of but two or three elements. But it is highly necessary to be able to deal with such simple types of combination.

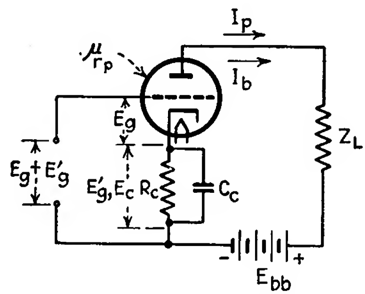


FIG. 143.—Circuit conditions of the grid-bias filter (R_c and C_c), which eliminates the necessity of the grid battery.

The Grid-bias Filter.—One of the most widely used combinations is the grid-bias filter, a resistance shunted by a capacitance. This combination is connected in series with the cathode of the tube and hence is in series with the grid circuit as well as with the dynamic plate resistance of the tube and the load impedance, as shown in Fig. 143.

The purpose of the grid-bias filter is to apply a direct voltage in series with the grid in order to maintain the grid at the desired negative potential. At the same time no appreciable *alternating* voltage is to be applied to the grid circuit by the filter, since this would interfere with the signal voltage to be amplified. The effect of the filter is, in other words, to select the direct-current part of the plate current, and to use it for grid-bias purposes, and at the same time to restrain the alternating-current

part of the plate current (namely, the amplified signal) from affecting the grid circuit. Nearly every radio-receiving tube now in service employs this method of obtaining the proper grid bias on the tube.

The explanation of the action of the grid-bias filter is as follows: The bias resistor R_c (Fig. 143) has a steady direct plate current I_b flowing through it owing to the action of the plate battery E_{bb}. As a consequence of Equation (55) it develops a direct voltage $E_{c_{\text{d.c.}}} = I_bR_c$ which is applied in the grid circuit, in the proper polarity to make the grid more negative than the cathode. At the same time, whenever a signal is present, the current in the resistor R_c contains one or more sine-wave alternating-current components, of amplitude I_p. These alternating-

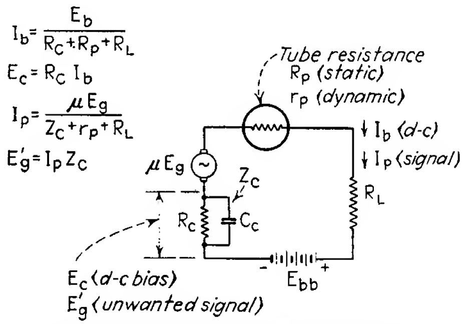


FIG. 144.—Equivalent circuit of the grid-bias filter (*cf.* Fig. 143).

current components produce a voltage across the resistor, but these are not wanted since they interfere with the weak grid voltage applied from the signal source. The condenser, having a low impedance to alternating current, "short-circuits" the alternating-voltage components, by constraining them to appear across the tube resistance and the load impedance. The amount of this action depends on the size of the condenser, the frequency of the signal current, and the tube and load impedances. By applying the impedance formulas for each unit, and by combining them according to the rules of Equations (64) and (65), it is possible to deduce the entire action of the plate circuit, both for direct current and for any frequency of alternating current.

Treatment of the Plate Circuit of a Tube Employing a Grid-bias Filter.—Consider the circuit shown in Fig. 144. The dynamic plate resistance is r_p ohms, the load impedance is assumed to be

a simple resistance, R_L ohms (although often it may be an inductance or a more complicated combination), the bias resistor is R_c ohms and the bias condenser is C_c farads. The signal applied to the plate circuit is a sine-wave of amplitude μE_g (see page 114). It is desired to find the resulting sine-wave plate current.

First we form the equivalent impedance of the grid-bias resistor and capacitor in parallel, then add this equivalent impedance to the other series impedances. The former process is performed as follows:

$$\frac{1}{Z_c} = \frac{1}{Z_1} + \frac{1}{Z_2} = \frac{1}{R_c} + \frac{1}{(-j/2\pi f C_c)}. \tag{65a}$$

This equation results from Equations (56), (63), and (65). It is solved for Z_c by ordinary algebraic manipulation, remembering that $j^2 = -1$, and that terms containing j as a factor are combined only with other terms containing j as a factor. The result is as follows:

$$\begin{aligned} Z_c &= \frac{R_c}{1 + 4\pi^2 f^2 C_c^2 R_c^2} - j\frac{2\pi f C_c R_c^2}{1 + 4\pi^2 f^2 C_c^2 R_c^2}, \\ &= R_e - jX_e \end{aligned} \tag{65b}$$

where R_e has been written to symbolize the first term (called the resistive term) and X_e the second (reactive) term. To obtain the total impedance Z_p of the plate circuit we add to Z_e the other two series units r_p and R_L, by Equation (64), and obtain:

$$Z_p = (r_p + R_L + R_e) + (-jX_e). \tag{65c}$$

The plate current for the alternating-current sine-wave signal μE_g is

$$I_p = \frac{\mu E_g}{Z_p},$$

or

$$I_p = \frac{\mu E_g}{(r_p + R_L + R_e) - jX_e}. \tag{65d}$$

In evaluating this equation numerically, the denominator is written as the square root of the sum of two squares:

$$I_p = \frac{\mu E_g}{\sqrt{(r_p + R_L + R_e)^2 + X_e^2}}. \tag{66}$$

It will be remembered that R_e and X_e are complicated expressions and that a numerical solution of Equation (66) involves considerable labor.

So far as the direct current is concerned, if we substitute $f = 0$ in the expression for Z_e above (65b), we obtain

$$Z_e = R_c.$$

That is, the grid filter acts simply as a resistance, so far as direct current is concerned. For any other frequency, however, the

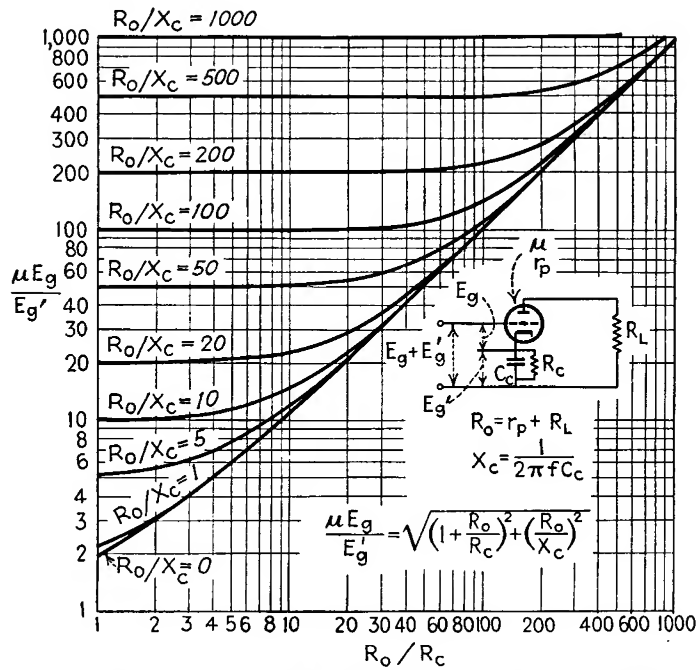


FIG. 145.—Ratio of unwanted grid signal E_g' to desired grid signal E_g, in terms of amplification factor, bias resistance, total plate-circuit resistance, and bias condenser reactance (μ, R_c, R_o, and X_c respectively).

voltage E_g' across the grid-bias filter is found by multiplying the alternating plate current I_p, from Equation (66) with the equivalent impedance of the grid-bias filter (65b), that is,

$$E_g' = I_p(R_e - jX_e) \tag{66a}$$

The phase relation between the current and voltage in the grid-bias filter is also important. This phase is given by

$$\phi = \tan^{-1}\frac{X_e}{R_e}. \tag{66b}$$

It is needless to say that this complete analysis serves for illustration only. In practice, the problem is solved simply by taking the value of R_c dictated by the direct-current component only and then selecting a capacitor such that its reactance [Eq. (63)] at the lowest frequency to be handled is not more than one-tenth to one-fifth as great as the resistance of R_c. This "rule of thumb" has been arrived at by process of experiment. Its actual efficacy can be gauged only by reference to the equations in the above analysis which, while complicated, give much insight into the actual working of the circuit. The reader is urged, therefore, to perform Prob. 1 at the end of this chapter.

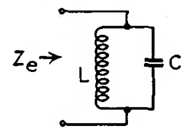


FIG. 146.—Ideal parallel-tuned circuit.

The Ideal Parallel-tuned Circuit.—The same process of combining the impedance values of various circuit parameters may be carried out in the highly important case of an inductance and a capacitance connected in parallel. This combination is called the ideal parallel-tuned circuit. It is a widely used illustration of the frequency-responsive characteristics of inductance coils and condensers. The case is much simpler than that of the complete plate circuit just considered, and leads to much more astonishing results.

Consider a pure inductance L (no resistance present) and a pure capacitance C connected in parallel as shown in Fig. 146. It is required to find the equivalent impedance of this combination. We use Equation (65) as before:

$$\begin{aligned}\frac{1}{Z_e} = \frac{1}{Z_1} + \frac{1}{Z_2} &= \frac{1}{j2\pi fL} + \frac{1}{(-j/2\pi fC)} \\ &= \frac{1}{j2\pi fL} + \frac{2\pi fC}{-j} \\ &= \frac{-j}{2\pi fL} + \frac{j4\pi^2 f^2 LC}{2\pi fL} \\ &= \frac{-j(1 - 4\pi^2 f^2 LC)}{2\pi fL}.\end{aligned}$$

And, inverting,

$$Z_e = \frac{j2\pi fL}{1 - 4\pi^2 f^2 LC}. \tag{67}$$

It will be seen that the equivalent impedance Z_e depends on the inductance and capacitance values and on the frequency, as we would expect. But there is much more important information present in the equation than this. If we plot the impedance Z_e against frequency f, for given values of L and C, we obtain the plot shown in Fig. 147. We find that at one particular frequency the impedance is enormously great, actually infinite if the inductance and capacitance are ideal, *i.e.*, not contaminated

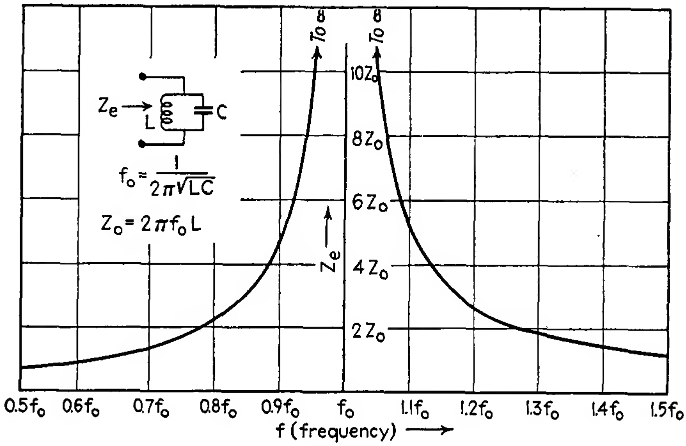


FIG. 147.—Equivalent impedance of the ideal parallel-tuned circuit, plotted against frequency of the applied voltage.

with resistance. The infinite impedance corresponds to the case when the denominator of Equation (67) is zero,

$$1 - 4\pi^2 f^2 LC = 0.$$

The frequency at which this occurs, obtained by solving the equation, is

$$f_r = \frac{1}{2\pi\sqrt{LC}}. \tag{68}$$

The frequency f_r is called the *resonant frequency*, because the circuit is particularly susceptible to ("resonates with") that frequency.

If a given alternating current of amplitude I is caused to flow through this parallel-tuned circuit, the voltage drop resulting across the circuit will have its maximum value when the fre-

quency of the current is equal to the resonant frequency, and it will have nearly as high values for frequencies slightly higher or lower than the resonant one. This ability of the parallel-tuned circuit to develop high voltage from currents within a restricted frequency range is widely employed in communications circuits, particularly in radio. In the latter case, the condenser in the tuned circuit is usually made variable, so that the frequency producing maximum voltage can be varied, and the circuit can be "tuned."

If an inductance and capacitance are connected in series, their combined impedance, by Equation (64), is

$$\begin{aligned} Z_e = Z_1 + Z_2 &= j2\pi fL - \frac{j}{2\pi fC} \\ &= \frac{j(4\pi^2 f^2 LC - 1)}{2\pi fC}. \end{aligned} \tag{69}$$

At the resonant frequency f_r given in Equation (68), the impedance of this series combination becomes zero. The series circuit thus possesses frequency-responsive properties, but the result is a heavy current (rather than a high voltage as in the parallel case) at the resonant frequency. Since vacuum tubes are, in general, voltage-operated devices, the parallel circuit is more commonly used than the series. Combinations of series- and parallel-tuned circuits, tuned to the same or to several different frequencies, can be made to reduce currents composed of many component frequencies into separate currents within restricted frequency ranges. These combinations, called "filters," are of inestimable importance in modern communications practice.

The presence of resistance in a tuned circuit, which cannot be avoided in practice since all wire has resistance, acts in general to lessen the degree of frequency responsiveness. Usually, however, the impedance of such a parallel-tuned circuit is very high at the resonant frequency, although not attaining infinity as in the ideal case. Similarly in a series-tuned circuit, at the resonant frequency, the impedance is not zero but is equal to the residual resistance value of the coil and condenser combination. The reader who is not familiar with tuned-circuit practice will benefit from the references given at the end of the chapter, which go into the subject much more in detail than is possible here.

54. Coupled Circuits.—All of the preceding circuit elements have been treated as "two-terminal" devices, that is, the impedance between two terminals only is considered. An important combination of circuit elements is represented by coupled circuits (Fig. 148), which are in general "four-terminal" devices, possessing two terminals to which an "input" current or voltage is applied, and two others from which an "output" current or voltage is obtained. The ordinary two-winding transformer, widely used in power practice and in communications, is an example of a coupled circuit.

The theory of the transformer is complicated since the reaction of the output circuit on the input circuit, and *vice versa*, must be taken into account. Ordinarily a simple and approximate theory can be used to describe the action of the transformer: The product of the input voltage times the input current is approximately equal to the product of the output current times the output voltage. The ratio of the input voltage to the output voltage is approximately equal to the ratio of the number of input turns to the number of output turns in the windings of the transformer. With these rules in mind it is possible to determine the output current and voltage for a given input current and voltage, and the reverse. A more complete analysis takes into account the power loss in the windings and the lack of complete magnetic coupling between the two windings. At high frequencies the capacitance between the windings and between turns of each single winding is also important. The effect of all these variables makes it possible to produce transformers which are sharply frequency-responsive or transformers which have little frequency discrimination. The latter type are important in communications circuits, in which it is desirable to transmit equally well many frequencies of alternating current.

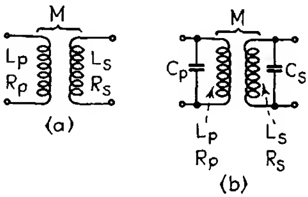


FIG. 148.—Untuned and tuned transformers (four-terminal impedance devices).

The tuned transformer, whose windings are commonly wound on a nonmagnetic core, has each winding shunted with a capacitor and hence consists of two parallel-tuned circuits magnetically coupled together. If the two circuits are tuned to the same resonant frequency, then a very great degree of fre-

quency selection is available. Furthermore, if the magnetic coupling between the coils is close, then the transformer will respond equally well to currents within a fairly broad range of frequencies and not appreciably to currents outside this range or "band." This type of transformer is called a "band-pass" transformer and is in effect a type of filter. It is much used in communications circuits, particularly radio-receiving circuits of the superheterodyne type.

55. The Electron Tube as a Circuit Element.—For most purposes it is sufficient to consider only two types of electron tube, the diode and the triode, as circuit elements. The tetrode

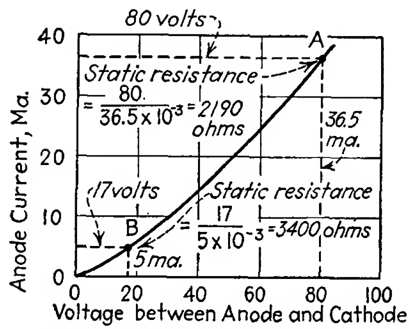


FIG. 149.—Graphical determination of the static plate resistance of a diode.

and pentode forms, while capable of a wider variety of circuit functions, are usually treated as equivalent triodes.

The Diode as a Circuit Element.—The vacuum-type diode has two characteristics of importance in circuit analysis, its static resistance and its dynamic resistance. The static resistance is the ratio of the total applied voltage to the total resulting current flow. This resistance value changes with the actual applied voltage, since the current increases with the $\frac{3}{2}$-power of the voltage. For example, in Fig. 149, the static resistance at an applied voltage of 80 volts is 2190 ohms, as indicated by the corresponding values of voltage and current at point A. At 20 volts, however, the static resistance has risen to 3400 ohms, as indicated by the point B. The value of static resistance, at the specified operating voltage, is used for direct-current calculations, *i.e.*, for determining the current flow when a given direct voltage is applied across the diode and a load resistor.

The use of the "load-line" method in this problem is described in a later section.

The dynamic resistance of the diode is the ratio of a small change in applied voltage to the corresponding small change in current which results from it. The value of the dynamic resistance depends on the point of operation of the tube. For example, referring to the graph in Fig. 57, (page 108) the dynamic resistance at 14 volts is 3200 ohms, at 72.5 volts 1666 ohms. This resistance value is used when dealing with the alternating-current component of current flow through the tube, provided that the amplitude of the alternating current is small compared with the direct current. If the entire current flow is alternating current, then the graphical method of Fig. 55 is used.

The Triode Tube as a Circuit Element.—The triode tube is a four-terminal device, the input terminals being the grid and cathode, the output terminals, the plate and the cathode. To each of these pairs of terminals are connected one or more circuit elements and one or more sources of voltage. The problem of circuit analysis in this case involves determining the voltages applied to the grid and the resulting current flowing in the plate circuit. (See Fig. 150.)

The grid circuit involves the grid impedance, across which the signal voltage is developed, and a source of direct bias voltage which acts in series with the signal voltage. The combined effects of these two grid-voltage components act to control the plate current. The plate current is caused to flow by the anode or "B" battery. In the absence of any signal only the static direct plate current flows, but when the signal is present, it causes an alternating-current component of plate current I_p to be superimposed on the direct plate current I_b. Thus in general there are two components to be considered in both grid and plate circuits, the bias or direct-current values and the signal or alternating-current values.

To show the relationship of these current and voltage values to the operating characteristics of the tube we can consider the general amplifier connection, shown in Fig. 150 and the static characteristics of the tube used, in Figs. 151 and 152. The direct bias voltage supplied to the grid circuit by the grid-battery is E_c volts. The signal voltage, supplied by the grid impedance Z_g is represented by a maximum amplitude E_g.

These voltages are shown measured along the e_g axis of the static characteristics (Fig. 151), the direct-current and alternating-current components adding as shown. The direct bias voltage applied to the plate is E_{bb}. Part (E_b) of this applied voltage appears across the tube resistance r_p, the remainder E_{bL} across the load impedance Z_L. The part of the direct voltage applied across the tube, E_b, with the value of the applied direct grid-bias voltage E_c, determines the static value of plate current I_b, as indicated on the static curves.

When a positive half cycle of the signal E_g is applied, it causes the grid voltage to become more positive, thus increasing the plate current. This increase in plate current causes an increase

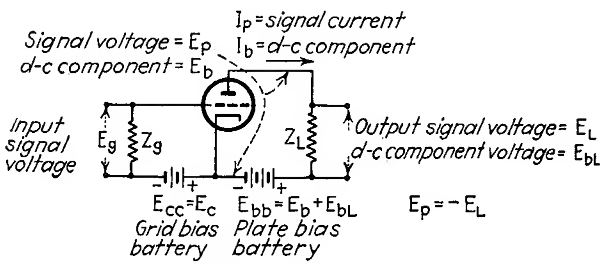


FIG. 150.—Current and voltage relationships in the basic triode amplifier connection. The letter symbols used are those listed in Appendix I, page 333.

in voltage drop across the load impedance. This additional voltage drop is subtracted from the applied battery voltage E_{bb} and the voltage across the tube E_b is reduced by that amount. The plate current thereby tends to be reduced somewhat in opposition to the increase caused by the grid-voltage change. When the grid signal voltage has reached the next (negative) half cycle, it makes the grid more negative, the plate current decreases, the load-impedance voltage drop decreases, the net voltage across the tube increases and the plate current tends to be increased somewhat, in opposition to the change produced by the grid-voltage change.

This description of the interrelation of the grid- and plate-circuit changes shows that when a signal is present, the net voltage applied to the grid, the net voltage applied to the plate, and the resulting plate current all change in unison. Consequently a considerable portion of the area of the static characteristics of the tube is occupied during the signal. As shown

in Fig. 151, the grid voltage ranges from $(E_c + E_g)$ to $(E_c - E_g)$, the plate voltage from $(E_b + E_L)$ to $(E_b - E_L)$, and the plate current resulting from these combinations of grid and plate voltage varies from $(I_b + I_p)$ to $(I_b - I_p)$. As the tube follows these ranges of voltage and current, it in effect traces out a

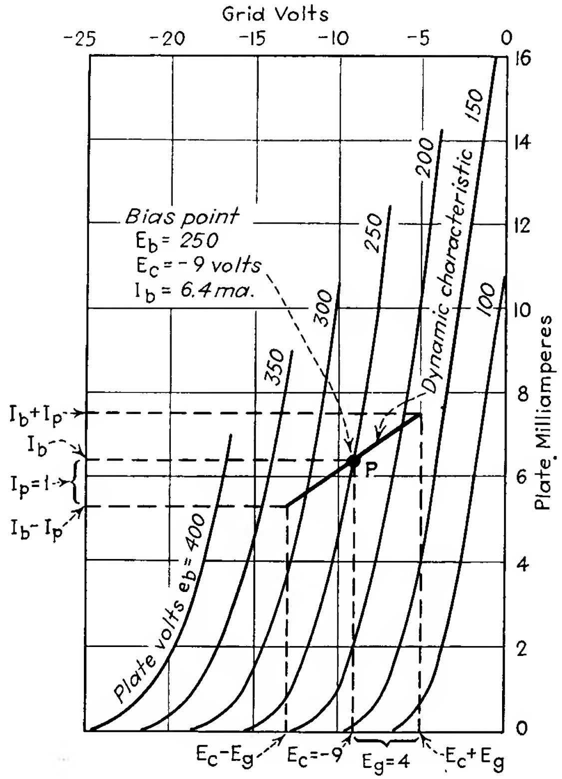


FIG. 151.—Dynamic characteristic of a triode tube working into a resistance load, plotted on the i_p-e_g curves. The letter symbols are those of Fig. 150.

locus on the static characteristics, as shown, which is called the *dynamic characteristic* of the circuit.

In the absence of signal the tube operates in the static condition represented by the bias point P, at which the net applied grid voltage and the net applied plate voltage (battery voltage less the voltage drop across the load) combine to produce the static plate current. When the signal is applied, the grid voltage covers the range shown, the plate voltage a corresponding range,

and the plate current a range resulting from the combination of the grid and plate voltages.

The Load-line Method of Determining Triode Performance.—If the load impedance is a simple resistance, the voltage drop across it is always proportional to the current flow through it and the current and voltage are always in phase. Hence it is possible to express the effect of the load resistance by a straight line, called the *load line,* whose slope is determined by the change

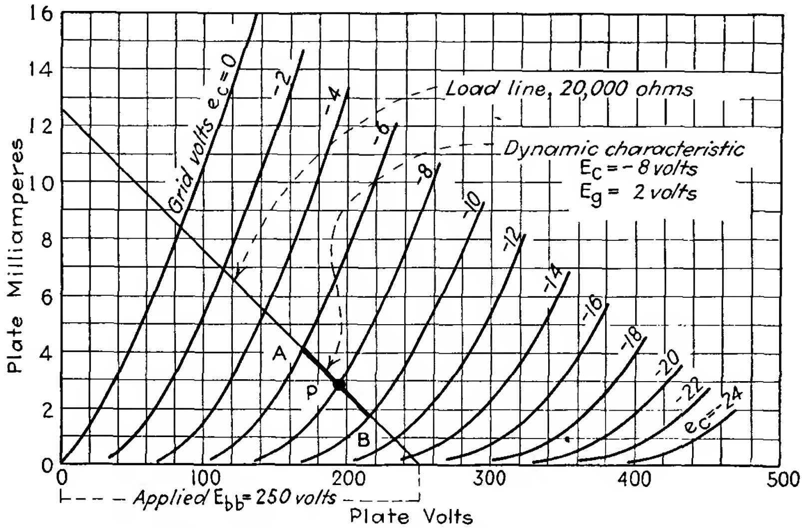


FIG. 152.—Load-line method of determining dynamic characteristics. The load line has an inverse slope equal to the load resistance value, and intersects the zero axis at the applied E_{bb} value. Its intersection with the curve corresponding to the applied grid-bias voltage marks the bias point P.

in current corresponding to a given change in voltage across the load, *i.e.*, the inverse of the resistance value of the load in ohms. This fact makes it possible to plot the dynamic characteristic of the circuit very readily, by means of the static characteristics of the tube and the load line. The static characteristics used are those involving as axes plate current and plate voltage, since the load-line slope is determined by these quantities. Such a family of curves is shown in Fig. 152. The bias point is given by the point P, determined by the applied grid-bias voltage and the net applied direct plate voltage (battery voltage less the direct-current drop in the load resistor). Through this point is

erected the load line, as shown. The slope of the line is such that a 1-ma. change in plate current corresponds to a change of 1 volt for every 1000 ohms present in the load resistance. This load line is the dynamic characteristic of the circuit, its end limits, A and B, being fixed by the limits of the change in the applied grid voltage. The intersection of the load line with the horizontal zero axis always occurs at a voltage value numerically equal to the battery voltage. The same system is, of course, applicable to diodes as well as triodes. In the diode the load line determines the plate current from a given applied voltage and load resistance.

If the load impedance is not a simple resistance, then a straight line no longer suffices to give the dynamic characteristic of the circuit. If, for example, the load impedance is a perfect tuned circuit, the current through it and the voltage across it are always 90° out of phase, and resulting dynamic characteristic must be an ellipse. These elliptical dynamic characteristics are useful in the study of the oscillator, in which one or more tuned circuits are always used. Such characteristics may be obtained experimentally with a cathode-ray-tube oscillograph.

Phase Reversal in an Amplifier Stage.—In certain applications, notably in television amplifiers and in push-pull amplifiers (see page 278), it is necessary to take into account the reversal of the phase of the output voltage of an amplifier with respect to the input voltage. Ordinarily, with a resistive load, the phase difference between these voltages is 180°, that is, the voltage across the load resistor becomes more negative as that across the grid resistor becomes more positive. The phase relationships between grid voltage, plate current, cathode-anode voltage, and output (load) voltage are shown in Fig. 153. To understand the phase reversal, consider an amplifier whose grid is biased sufficiently negative to cut off the plate current completely. Since no current then flows through the load resistor, there is no voltage drop across it, and the two ends of the resistor are at the same potential. The upper end of the load resistor, like the lower end, has a high positive potential with respect to the cathode, because of the presence of the battery E_{bb}. Now we make the grid voltage more positive, thus allowing plate current to flow through the load resistor. The voltage drop thereby appearing across the load resistor is subtracted from the plate

battery voltage, E_{bb}, and the upper end of the load resistor assumes a less positive (more negative) potential, while the upper end of the grid resistor becomes more positive. Consequently the output load voltage is reversed in phase with respect to the grid voltage. The voltage between anode and cathode of the tube is, on the other hand, in phase with the grid voltage, since the sum of the cathode-anode voltage and the output

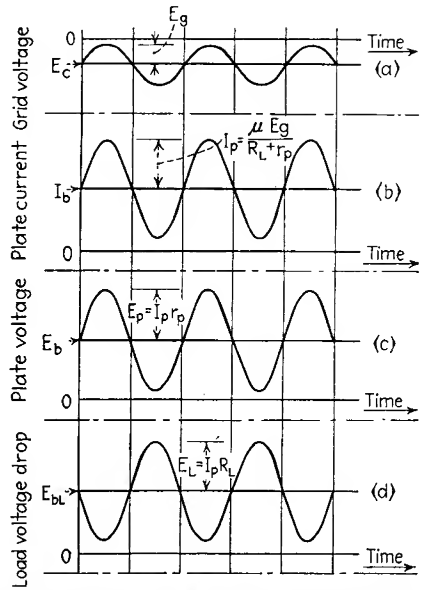


FIG. 153.—Phase relations in the grid and plate circuits of a triode working into a resistive load. The grid voltage, plate current, and cathode-anode voltage are in phase. The cathode-anode voltage, subtracting from the fixed battery voltage, produces a load voltage 180° out of phase with the applied grid voltage. (Letter symbols are those of Fig. 150.)

voltage must at all times equal the applied battery voltage, that is, $E_b + E_{bL} = E_{bb}$, and $E_p = -E_L$ (*cf.* Figs. 150 and 153).

56. Power Relationships in Electronic Circuits.—Power, the rate at which energy is expended, is measured by the instantaneous product of the current and voltage in a circuit. In direct-current circuits, in which the current and voltage do not vary, the power is a constant quantity. In alternating-current circuits, however, the power varies from instant to instant, since it depends on the instantaneous product of *alternating* current and voltage. Usually the changing power in alternating-

current-circuits level is averaged over one or more cycles of the alternating current, and this average power value is used to determine the over-all requirements of the equipment. The calculation of the average power must take into account the phase angle between the current and voltage. For example, the six diagrams shown in Fig. 154 show a given current and voltage in phase, 90° out of phase, and 45° out of phase, respectively, together with the curves which represent the instantaneous product of the current and voltage, *i.e.*, the instantaneous power

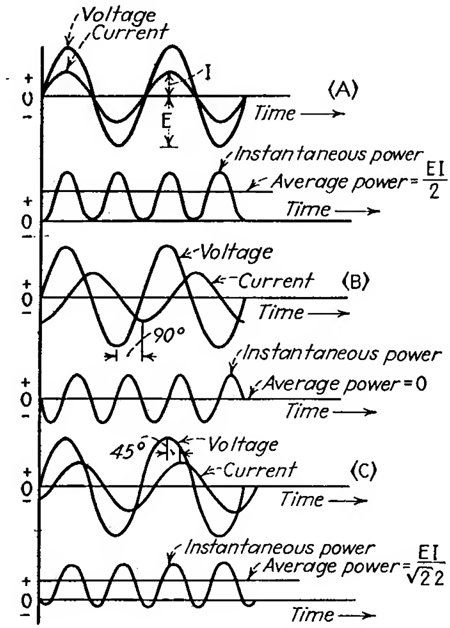


FIG. 154.—Effect of phase angle on power.

at each point in the cycle. If the average of the first case is 1 watt, then the average of the second case is zero, and the average of the third case is $1/\sqrt{2}$ watt, the difference being caused simply by the variation in phase between current and voltage.

The instantaneous power delivered to or from a circuit element carrying an alternating current of I amp. ($I =$ maximum amplitude) when the applied alternating voltage is E volts, maximum, and the phase difference is ϕ, is

$$P_{av} = \frac{EI}{2} \cos \phi \text{ watts.} \tag{70}$$

This relation can be used to compute the average power delivered by a triode to a load resistance, using the load-line method. The load line, found as in Fig. 152, gives the total range of current and voltage across this load resistance. The total range in each case is twice the amplitude of the alternating voltage and current, *i.e.*, $(E_{\text{max.}} - E_{\text{min.}}) = 2E$, and

$$(I_{\text{max.}} - E_{\text{min.}}) = 2I,$$

where E and I are the maximum amplitudes of the alternating current and voltage across the load. Substituting these values in Equation (70) and remembering that the phase angle is zero ($\cos 0 = 1$), we obtain

$$P_{\text{av}} = \frac{(E_{\text{max.}} - E_{\text{min.}}) \times (I_{\text{max.}} - I_{\text{min.}})}{8}. \tag{71}$$

R-m-s Values of Current and Voltage.—The ordinary alternating-current voltmeter and ammeter indicate not maximum values but r-m-s or "effective" values which are 70.7 per cent as great as the maximum value of the sine wave. These effective values are equal numerically to the values of a direct current and voltage which will produce the same power in a given resistance. In a resistance, in which there is no phase difference, Equation (70) becomes

$$P = \frac{EI}{2}, \tag{72}$$

whereas the power delivered by a direct current and voltage is

$$P = E_{\text{d.c.}} \times I_{\text{d.c.}}. \tag{73}$$

Using the r-m-s values, the power in an alternating-current circuit is

$$P = E_{\text{r.m.s.}} \times I_{\text{r.m.s.}}. \tag{74}$$

Since Equations (72) and (74) express the same thing, it follows that

$$E_{\text{r.m.s.}} = \sqrt{2}E = 0.7071E \tag{75}$$

and

$$I_{\text{r.m.s.}} = \sqrt{2}I = 0.7071I. \tag{76}$$

In all relations up to Equation (70) maximum amplitudes (E and I) have been used. If r-m-s values are substituted in these relations, they must be converted to maximum amplitude values by Equation (75) or (76) before substitution.

Problems

1. In the circuit shown (Fig. 143), a sine-wave alternating-current signal, frequency 100 cycles and amplitude 0.5 volt, is applied between the grid and cathode. Find the amplitude of the sine-wave alternating-current component of the plate current. Find the resulting alternating voltage drop across the grid-bias filter. $R_c = 1000$ ohms; $C_c = 2.0\mu f$; $R_L = 10,000$ ohms; $\mu = 20$; $r_p = 10,000$ ohms.

2. Compute the impedance of an ideal parallel-tuned circuit, $C = 1.0\ \mu f$, $L = 1.0$ henry, at 150 cycles, 160 cycles, and 170 cycles, and the voltage drop across the circuit in each case if the maximum amplitude of the current flow is 0.1 ma.

3. The diode of Fig. 54 is connected to a 2,000-ohm load resistance and to a mixed voltage source composed of 10 volts (maximum amplitude) 60-cycle alternating current superimposed on 100 volts direct current. Compute the corresponding direct- and alternating-current flow. Repeat the computation for 30 volts alternating current superimposed on 300 volts direct current. Use a load line for determining the direct voltage applied across cathode and anode.

4. A triode (static characteristics in Fig. 152) is connected to a load resistance of 15,000 ohms. The grid-bias battery is -4.0 volts, while the plate battery is 150 volts. Find the bias point on the static characteristics.

5. The grid circuit of the tube in Prob. 4 is fed a sine-wave alternating-current signal of 2.0 volts maximum amplitude. Draw the dynamic characteristic of the circuit, and compute the average alternating-current power fed to the load impedance by the tube.

6. Compute the total (alternating-current plus direct-current) average power supplied by the plate battery in Prob. 5 and the efficiency with which this power is transferred as useful alternating current to the load impedance.

7. Repeat Probs. 5 and 6 if the plate battery voltage is 300 volts. Why is the efficiency lower in this case?

Bibliography

TIMBIE and BUSH: "Principles of Electrical Engineering," John Wiley & Sons, Inc., New York, 1930.

LAWRENCE, R. R.: "Principles of Alternating Currents," McGraw-Hill Book Company, Inc., New York, 1935.

WEINBACH, M. P.: "Alternating Current Circuits," The Macmillan Company, New York, 1933.

GUILLEMIN, E. A.: "Communication Networks," vol. I, John Wiley & Sons, Inc., New York, 1931.

EVERITT, W. L.: "Communication Engineering," McGraw-Hill Book Company, Inc., New York, 1937.

Albert, A. L.: "Electrical Communication," John Wiley & Sons, Inc., New York, 1934.

Henney, Keith: "Principles of Radio," John Wiley & Sons, Inc., New York, 1938.

Terman, F. E.: "Radio Engineering," McGraw-Hill Book Company, Inc., New York, 1937.

Glasgow, R. S.: "Principles of Radio Engineering," McGraw-Hill Book Company, Inc., New York, 1936.

Chaffee, E. L.: "Theory of Thermionic Vacuum Tubes," McGraw-Hill Book Company, Inc., New York, 1933.

CHAPTER XII

POWER-TRANSFORMATION CIRCUITS

Introduction.—By power transformation is meant the modification of the amplitude, frequency, or phase of a basic power source. Amplitude transformation, changing a high-current, low-voltage source into a high-voltage, low-current source, and *vice versa,* is commonly carried out in alternating-current practice by transformers; in direct-current practice, electron tubes are called into play. Frequency transformation includes rectification (as well as the reverse process, called inversion), and frequency conversion, changing a low-frequency source into a high-frequency one and *vice versa.* These processes are evidently of great importance in all branches of applied electronics. In this chapter we discuss their application to basic power sources, rather than to the communications and control problems which are discussed in the following chapters.

57. Rectification: Half Wave, Full Wave, Multiphase.—When an alternating-current source is to be converted electronically into a direct-current source, rectifier tubes in one form or other are used. These tubes may be vacuum- or gas-filled, they may be simple diodes or triodes, but their essential function is one-way conduction, the passage of electrons from cathode to anode—not in the reverse direction. Various forms of circuit which have been developed for utilizing this function differ primarily in their economy of operation and in the ease with which they may be controlled.

The simplest circuit is the half-wave rectifier, shown in Fig. 155. The alternating-current power source is connected in series with the load and the rectifier tube. When the polarity of the alternating current makes the anode positive, conduction occurs; on the remaining half cycles the tube does not conduct. The current through the load, plotted against time, then has the form shown in the figure. This is not a direct current, rather it is a unidirectional one containing both alternating-current and

direct-current components. The alternating-current part may be converted into a steady direct-current flow by smoothing circuits or filters, a typical form of which is shown in Fig. 156. Smoothing filters of this type consist of series inductance and shunt capacitance, the resonant frequency of the combination being considerably lower than the frequency of the alternating-current supply. Figure 157 gives operating characteristics of typical filter sections. The average current carried by a half rectifier is 0.318 times the peak current. The main disadvantages of the circuit are: (1) The current in the load contains an alternating-current component whose basic frequency is equal to the power frequency, so that elaborate filtering is required to produce smooth direct current, and (2), the alternating-current power supply is delivering power only half the time, and its ratio of peak load to average load is therefore high. Both these objections are answered in part by the full-wave and multiphase rectifiers described below.

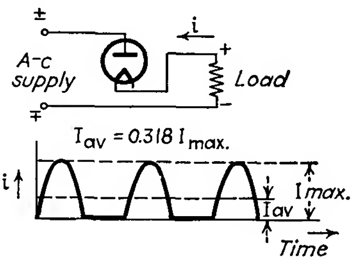


FIG. 155.—Basic diagram and current wave form of the half-wave rectifier.

In half-wave rectifier circuits diodes are usually used, although gas-filled triodes are used in half-wave controlled-rectifier

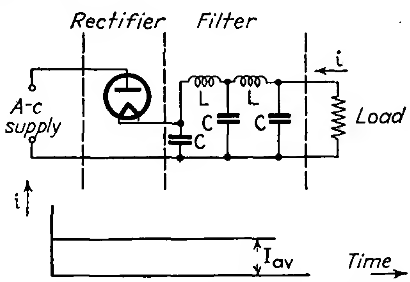


FIG. 156.—Effect of smoothing filter on current wave form.

circuits (page 259). Vacuum diodes are used when the current requirements are not great and when high efficiency is not essential. For the heavy-current, high-efficiency applications, gas-filled diodes are universally used.

The Full-wave Rectifier.—The circuit shown in Fig. 158 contains two rectifier tubes which work on alternate half cycles, the power

supply connections being such that the anode of one tube is positive while that of the other is negative, and *vice versa.* The load is placed in the common return of the two cathodes and thus receives the current flow from both tubes. As a con-

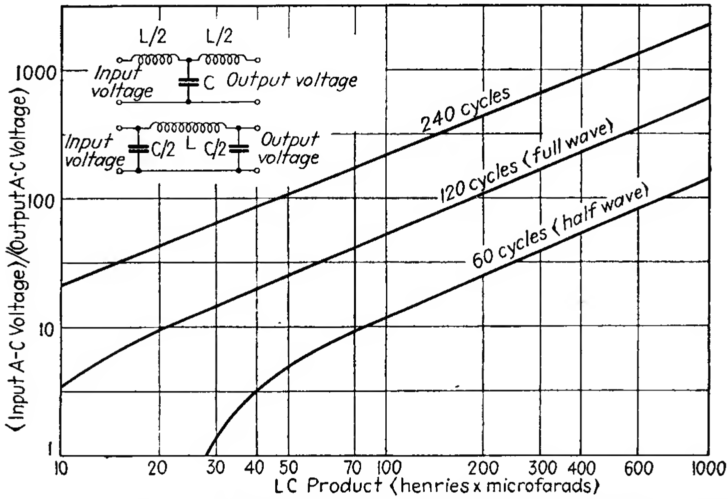


FIG. 157.—Reduction of alternating-current component in terms of inductance and capacitance values in the filter. (*Data from W. W. Waltz.*)

sequence, the current in the load, plotted against time, has the form shown in the figure. It will be noticed that this pulsating unidirectional current has an alternating-current component whose basic frequency is twice that of the power supply. The smoothing filter may then have L and C values only 50 per cent

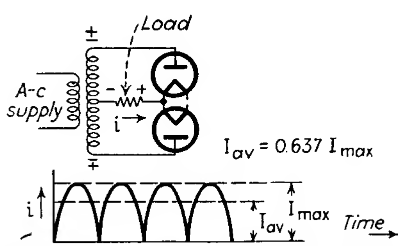


FIG. 158.—Connection and wave form of the full-wave rectifier.

as great as are needed in the half-wave rectifier. The power supply delivers current equally on both halves of each cycle and hence is used more efficiently, at least so far as the primary winding of the transformer is concerned. Each half of the

center-tapped secondary of the transformer, however, carries current only on alternate half cycles. The full-wave rectifier is used in most alternating-current-operated radio receivers, and in similar medium-power applications.

The need of the center-tapped supply transformer is eliminated in the bridge-type rectifier, another full-wave type, in which four tubes are used. The circuit, shown in Fig. 159, is so arranged that two tubes carry the load current in series, on alternate half cycles. The utilization of the power supply is thus at a maximum. However, the use of four tubes is somewhat uneconomical. For high-voltage work, the bridge rectifier is popular since the inverse (nonconducting) voltage is applied in series across two tubes, which therefore divide the strain between themselves.

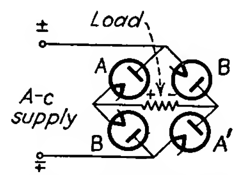


FIG. 159.—Bridge-type full-wave rectifier.

The Voltage-doubling Circuit.—The circuit shown in Fig. 160, called a *voltage doubler*, contains two separate half-wave rectifiers. Each tube charges its associated condenser on alternate half cycles, and the condensers are so connected that the voltage of one adds to the voltage of the other. Such a circuit, at no load, will produce about 300 volts from a 110-volt r-m-s supply. When a filter and load are connected, however, the voltage drop in the rectifier tubes and the discharge of the condensers are enough to lower the output voltage considerably. In this circuit the two cathodes must be isolated electrically, so that double rectifiers having a common cathode (used in full-wave circuits) cannot be used.

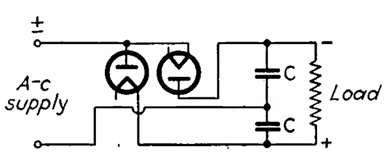


FIG. 160.—Voltage-doubler circuit.

Multiphase Connections.—In alternating-current power practice it is usual to supply power over three wires, in a so-called three-phase circuit. The voltage appearing between any two wires is 120° out of phase with those appearing across the other two pairs; each wire serves to some extent, therefore, as the return wire for the other two, and the efficiency of transmission (and utilization) of the power is thereby raised by raising the average level of the current carried, relative to the peak value.

When such a three-phase system is available, it is customary to take advantage of its high efficiency by connecting rectifier tubes in three-, six-, or twelve-phase circuits. Two typical six-tube circuits are shown in Figs. 161 and 162. Each tube carries current only during the positive-anode half cycles of the alternating current presented to it, but the load is connected so

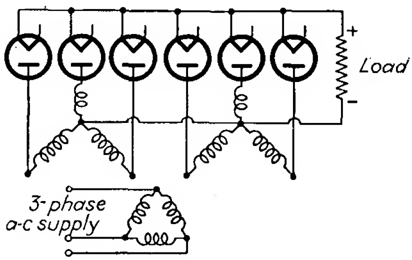


FIG. 161.—Double-Y three-phase rectifier.

that the cathode currents of all the tubes flow through it, and as a consequence the average load current is much higher relative to the peak current than in the single-phase rectifiers. The alternating-current component of the output current is thus small in magnitude and its frequency is three or six times as great as the supply frequency. Consequently relatively small

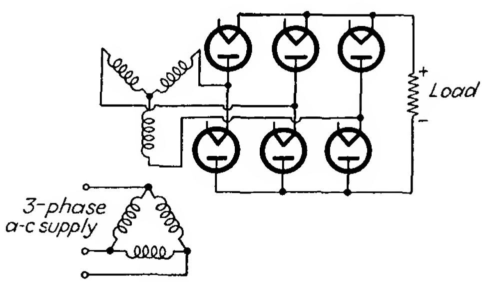


FIG. 162.—Three-phase full-wave rectifier.

filter components will suffice to produce smooth direct current from the pulsating current supplied by the rectifier.

Multiphase rectification is used only when efficiency is all-important, *i.e.*, when the cost of the power handled is commensurate with the cost of the rectifier itself. Since efficiency is thus paramount, the tubes used are almost universally of the high-efficiency gas-filled type. The power supply for broadcast

stations and other apparatus requiring up to 20,000 volts direct current at peak currents of 10 amp., makes use of multiphase gas-filled rectifiers.

58. Grid-controlled Rectifier Circuits.—In all the rectifier circuits thus far considered, the power delivered by them to a given load depends solely on the supply voltage. If more or less power is to be delivered to the load, the change can be made by changing the supply voltage, by the insertion of a variable transformer, or a series resistance or inductive reactance. All these methods can be, and are, used when a controlled power level is required. But they are wasteful of power, since some power is usually consumed in the control of the supply voltage, and the control equipment, since it must carry the full-load current and voltage, is costly and cumbersome to operate.

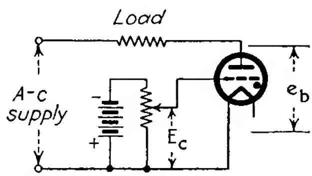


FIG. 163.—Grid-controlled gas-triode rectifier.

A completely different approach to the problem is made available by the grid-controlled tube. Either gas-filled or vacuum triode tubes may be used, but since economy is usually the first consideration, gas-filled tubes are used to the exclusion of the vacuum types. The principle used in the gas-filled controlled rectifier is the control of power by controlling the time during which the current flows in each cycle. In brief, the grid of the gas triode is used to withhold the start of the current flow until some predetermined part of the cycle. After the start of the discharge, of course, the current flow continues until the end of the half cycle, since the grid is then powerless to control the discharge.

A simple form of the controlled rectifier is shown as Fig. 163. It is a half-wave circuit in which an adjustable direct voltage is applied between grid and cathode. The power delivered by this circuit depends on the time in the conducting cycle at which the discharge starts, and this in turn depends on the control characteristic (Sec. 33, page 147) of the tube used. The control characteristic as ordinarily plotted does not serve to reveal the desired information, so it is customary to use the control characteristic as the basis of another plot, called the control-cycle diagram, an example of which is shown in Fig. 164. Above the

time axis is plotted one half cycle of the anode voltage actually applied to the tube. Below the time axis are plotted values of grid voltage, the value plotted at each point being that grid

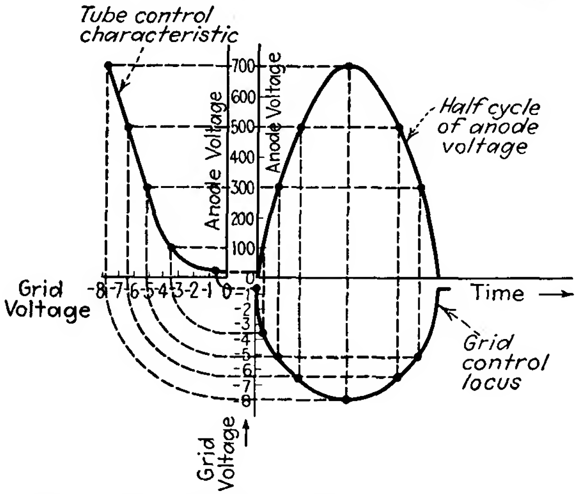


FIG. 164.—Graphical determination of grid-control locus from applied plate voltage and tube-control characteristic.

voltage which will just cause conduction to start when the anode voltage has the value directly above that point, as indicated in the diagram. The grid-control locus so plotted serves to show at what point in the cycle the applied grid voltage will start the discharge. For example, if the adjustable grid voltage (E_c, Fig. 163) is given the value -2.0 volts, this applied voltage may be represented on the control-cycle diagram by a straight line, displaced 2.0 volts below the time axis. The point at which this straight line intersects the grid-control axis (see Fig. 165) is the point at which both anode and grid voltage have the proper combination of values to start conduction. Accordingly, at the time corresponding to this intersection the conduction starts, and continues thereafter until the end of the half cycle.

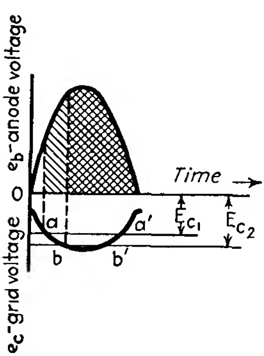


FIG. 165.—Simple direct-current amplitude control of gas-triode rectifier.

By varying the applied grid voltage (shifting the position of the horizontal line up or down), the point of intersection can be made to occur at any point from the beginning of the cycle to the mid-point. If, however, the grid voltage is made more negative than this, no intersection occurs at all, the condition for firing is not satisfied, and the tube remains nonconducting throughout the entire cycle. This simple method thus permits control from full current to one-half current, but no control is provided from one-half current to zero current. Furthermore, as is clear from the diagram, any slight relative displacement of the applied voltage line and the grid locus will cause a large

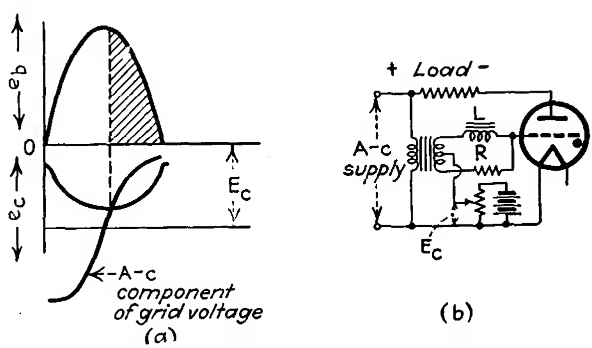


FIG. 166.—Alternating-current plus direct-current amplitude control.

shift of the time of intersection. Since such a displacement can be caused by variations either in the applied grid voltage or in the tube-control characteristic, erratic operation is apt to result, especially when the intersection is near the mid-point of the cycle.

Alternating-current Control of Gas-triode Rectification.—These objections (limited control range and erratic operation) have led to the development of the alternating-current type of grid control, in which a combination of alternating and direct current (or in some cases alternating current alone) is applied for control purposes. The applied alternating current is the same frequency as the applied anode supply, but displaced from it in phase by 90°. The position of the grid voltage, with respect to the anode voltage, is as shown in Fig. 166. The intersection of the grid voltage with the control locus is then very sharp. Furthermore, by varying the direct voltage applied to the grid (or when

alternating current alone is used, by varying the amplitude of the alternating current) the point of intersection with the grid locus may be changed from the beginning of the cycle to a point considerably beyond the mid-point. The range and reliability of control are thus considerably extended beyond those attainable by simple direct-current control.

Even the variable-amplitude alternating-current system has a somewhat limited control range, since the intersection point cannot be displaced completely to the end of the cycle. The most flexible system (and most widely used) is phase-shift control, in which the phase displacement between grid and anode voltage is controlled by a variable-reactance combination

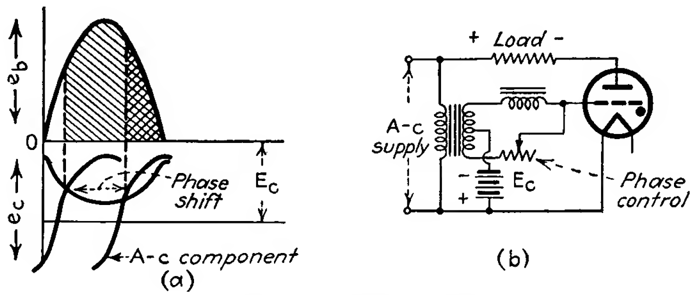


FIG. 167.—Phase-shift control.

in the grid circuit, as is shown in Fig. 167. The phase shift is controlled by varying the resistance value. The amplitude of the direct-current grid bias may be adjustable, but ordinarily it remains at a fixed value. As the phase of the grid voltage is shifted, the point of intersection with the grid-control locus is advanced or retarded, as shown, and the range extends from the beginning of the half cycle to the end. The power control is thus continuous from full output to zero. If a three-phase system is available, a regulation phase-shifting transformer may be used in the grid circuit, in place of the *R-L* combination shown in the figure.

The uses of these grid-controlled rectifier units (see Chap. XIV) range from the dimming of theater lights to the automatic regulation of alternating-voltage supplies. The control mechanism in each case is small, inexpensive, and simple to maintain, and the control of the power is continuous and smooth. One theater-light installation (that in Radio City, New York)

contains 328 separate circuits, all controlled from a master dimming panel. A typical "feedback" circuit used in this installation is shown in Fig. 168.

59. Inversion Circuits.—The production of alternating current from a direct-current supply is called inversion. Two types of circuits are used for this purpose, the vacuum-tube oscillators and the gas-filled triode inverters. The oscillators are comparatively inefficient and are usually used only where the alternating current produced must be of higher frequency than the gas-filled tube can handle. For frequencies up to approxi-

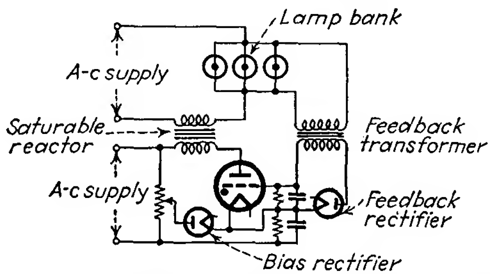


FIG. 168.—Feedback lamp-dimmer circuit. The bias on the grid-controlled rectifier is obtained in part from the dimming control, and in part from the lamps themselves.

mately 5000 cycles, the gas-filled-tube inversion circuits are used.

Single-tube Inverter Circuits.—The simplest inverter circuit contains but one tube, which acts as a switch. As shown in Fig. 169, the tube is connected in series with the primary of a transformer, which thereby conducts pulses of current as the tube becomes alternately conducting and nonconducting. These direct-current pulses in the primary of the transformer become alternating-current pulses in the secondary. The system is exactly analogous to the ignition system of an automobile, in which direct current from the generator is interrupted periodically by a mechanical contactor in series with the primary of the ignition coil.

The periodic interruption of the current flowing through the gas triode is accomplished by the use of a capacitor and resistor combination in the grid circuit. When the switch is closed (Fig. 169), the capacitor C charges from the direct-current source through the resistor R. Initially the voltage across the

capacitor is zero, but as time goes on, the voltage across the capacitor increases until at the end of $R \times C$ microseconds, it is 63 per cent of the full line value. This increasing capacitor voltage is applied across the transformer and the anode-cathode circuit of the tube, which is restrained from conducting current by the direct bias voltage applied to the grid of the triode. This bias, obtained by a voltage divider across the direct-current line, is of such a value that it allows conduction when the condenser voltage has risen to a value somewhat below the line voltage. At this point the tube conducts and begins to discharge the capacitor, causing a pulse of current to pass through the transformer primary.

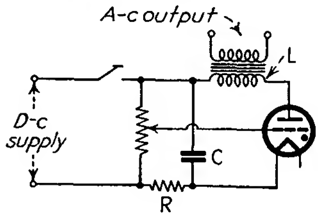


FIG. 169.—One-tube discharge inverter circuit.

The resistance of the tube and transformer primary is lower than the charging resistor R, hence the capacitor loses charge through the tube faster than it gains it from the supply, and the capacitor voltage falls more or less rapidly to a value low enough to extinguish the discharge in the tube. However, it is important that the tube continue conducting at this point, else the capacitor would immediately start recharging from the supply line and the tube would again conduct before deionization had taken place, *i.e.*, before the grid had regained control. So the inductance of the transformer primary is utilized to maintain the discharge for a short time after the extinction voltage level is reached. When the discharge finally ceases, therefore, the capacitor must consume some time in recharging to the ignition point, and sufficient time for deionization is thereby obtained. The tube then remains nonconducting until the capacitor has charged up to a value somewhat below full line voltage, at which the grid normally allows conduction to start. The capacitor thus successively charges from the line, discharges through the tube and transformer, recharges from the line, redischarges into the tube and transformer, and so on.

The power output of this type of inverter depends on the applied direct voltage and the current-carrying ability of the gas triode used. The frequency depends on the values of R and C used and on the ratio of the charging resistance R to

the discharging resistance (tube resistance and transformer-primary resistance in series). Since the transformer characteristics displayed by the primary depend to some extent on the load connected to the secondary, the frequency may be changed by changes in load, an undesirable feature. Furthermore, the wave form of the alternating current produced is far from a sine wave. The current in the primary is in the form of more or less discontinuous pulses, and the alternating current produced in the secondary has a correspondingly rough "saw-tooth" form.

Two-tube Inverter Circuits.—Two-tube inverter circuits are preferred for their greater time-average economy. In addition, the two-tube inverters are more stable and can be controlled from an external frequency-determining source. There are two important types of double-tube circuits, the parallel and the series.

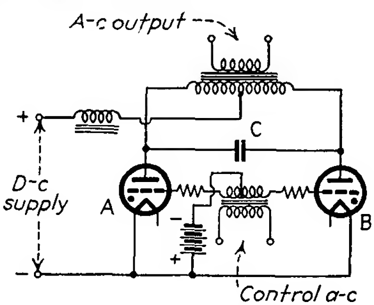


FIG. 170.—Full-wave parallel inverter.

A typical parallel circuit is shown in Fig. 170. The frequency of the circuit is controlled from an external source, through a grid transformer whose center-tapped secondary feeds alternating-current pulses alternately to the two grids, the pulses being superimposed on the required direct bias voltage. The anode currents of the two tubes feed into opposed sections of the primary of the output transformer, through whose center tap the direct voltage is applied to the anodes. The secondary winding of the output transformer supplies the generated alternating current to the load.

In this inverter, as in any gas-filled-tube device, some means must be provided for reducing the anode voltage of each tube below the extinction point once for each cycle, so that the grid of each tube regains control of the discharge in proper succession. In the parallel inverter the reduction of anode voltage is obtained by the connection of the capacitor C (Fig. 170) between the anodes. To illustrate the action, consider the alternating control voltage applied to the grids at the proper instant to make the grid of tube A positive. Tube A thereupon conducts

and will remain conducting until its anode voltage is reduced to the extinction point. When the control alternating voltage reaches the next half cycle, the grid of tube *B* is made positive and this tube conducts. When it conducts, the voltage between its anode and cathode is sharply reduced from the full direct line voltage (say 115 volts) to the normal tube voltage drop, which is 10 or 15 volts. This sudden decrease in voltage (approximately 100 volts) across tube *B* is transferred, through the capacitor *C* (a capacitor has no initial impedance to a sudden pulse or change of voltage), to the anode of tube *A*. The voltage drop across tube *A* (then conducting) is 10 or 15 volts, but when the sudden reduction of voltage is transferred to it, its voltage actually becomes 85 volts *negative* with respect to its cathode, at least momentarily. This momentary reduction of anode voltage extinguishes the arc in tube *A* and allows time for deionization to occur. Thereupon the grid of tube *A* has control and does not allow conduction until the polarity of the control alternating voltage again makes it positive. When tube *A* then starts, the sudden drop of voltage across it is transferred to tube *B*, which is thereby extinguished. The conduction thus passes back and forth from tube to tube, each conduction period sending pulses through the output transformer primary, which generates the required alternating current in the secondary.

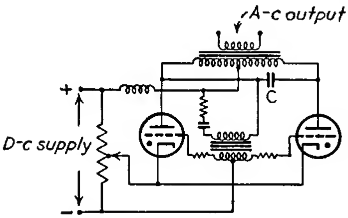


FIG. 171.—Self-excited full-wave inverter.

If the control voltage on the grids of the tubes is obtained not from an external alternating-current source but from the output transformer itself, then the inverter is said to be self-excited; it produces alternating current without reference to any other source of alternating current. A self-excited inverter circuit of the parallel type is shown in Fig. 171. The frequency stability of this type is not so great as that of the externally synchronized type, but it is considerably better than the single-tube inverter. The frequency of the self-excited inverter is determined primarily by the *R*, *L*, and *C* constants in the grid circuit and the inductance exhibited by the output transformer primary. The power output depends on the applied direct voltage and the current

capacity of the tubes. Note that by adjusting the direct bias voltage the duration of the conduction in each half cycle can be controlled, *i.e.*, the tube then works as a grid-controlled rectifier in reverse, and the power output (amplitude of the output alternating voltage and current) as well as the frequency is under control.

The Series-type Inverter.—One form of series-inverter circuit is shown in Fig. 172. The capacitor C is charged through tube A, then discharged through tube B, as the grids of the two tubes are excited from the external alternating-current source. When the discharge current through tube B starts, a sudden change in voltage is built up across the inductance (choke coil) L_2 and this voltage raises the potential of the cathode of tube A, extinguishing it and allowing deionization to take place. When tube A subsequently conducts, a similar change in voltage across the choke L_1 extinguishes tube B.

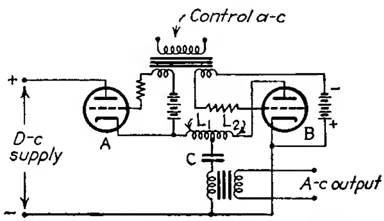


FIG. 172.—Series full-wave inverter.

Oscillator Inverter Circuits.—Direct current may be transformed into alternating current by means of vacuum-tube oscillator circuits, which are essentially amplifiers containing a tuned circuit in grid or plate circuits, or both, which establishes the frequency of the generated alternating current, and a connection of some sort between grid and plate circuits, which establishes the self-excitation feature already noted. These oscillator circuits are applied much more widely to communications practice than to basic power sources, and hence discussion of them is deferred to Chap. XIII. They are distinguished by purity of wave form of their output power, efficiency usually not greater than 70 per cent (as against 90 per cent for inverter circuits), and the need of large tubes and high voltage to obtain high power output.

Frequency-multiplication and -division Circuits.—Electronic circuits are used not only to convert alternating into direct current and *vice versa* but also to change the frequency of alternating current from one value to another. This may be done in oscillator or inverter circuits by arranging the natural resonant

frequency of the output to be some multiple of the input control frequency. The output circuit is then energized every other cycle, every third cycle, or in extreme cases every fourth cycle, giving rise to frequency doubling, tripling, and quadrupling, respectively. The need for this sort of power transformation is not great, so frequency-multiplication circuits have not been developed extensively for power practice. They are widely used in radio-frequency applications, as is shown in the following chapter.

The inverse process, frequency division, is not so easily accomplished. It is not sufficient to make the resonant frequency of the output some submultiple of the excitation frequency, since the output would then be energized twice, three times, or four times per cycle, etc., and the net result would be no output at all or else an output equal to the excitation frequency. Frequency division can be accomplished in relaxation circuits, called multivibrators, in which the output circuit is energized through a capacitor which stores several pulses before delivering them (see Chap. XIII).

60. Rectifier-inverter Combinations; the Direct-current Transformer and the Direct-current Transmission of Power.— The use of rectifier and inverter circuits in combination makes it possible to transform alternating into direct current and back again as often as is required. This makes possible the so-called "direct-current transformer" which converts a high-current low-voltage direct-current source into a high-voltage low-current direct-current source, or *vice versa*. The input direct current is first converted into alternating current at some convenient frequency, which is then fed to a step-up or step-down transformer (depending on whether an increased or decreased direct-voltage output is required). The output alternating current of the transformer is then rectified and filtered, *i.e.*, converted back to direct current. Exactly this process is used in every automobile radio receiver, in which high voltage for the plates of the tubes must be obtained from the 6-volt-storage-battery supply in the car. In this case a nonelectronic mechanical inverter (called a vibrator) is used to produce the alternating current from the direct-current input. The rectification usually is carried out by a vacuum-tube full-wave rectifier.

The other inverter-rectifier combination of importance is used in the direct-current transmission of power over high-voltage

transmission lines. Power is ordinarily sent over long distances in alternating-current form, since alternating current can be generated at high voltage more economically than direct current. But there are many drawbacks: Alternating current tends to flow on the surface of the transmission wires, and hence the whole cross-section area of the wire is not utilized. The insulation must be designed to handle the peak values of the voltage, which are 40 per cent higher than the effective r-m-s values and exist only for a short time. Power-factor correction must be provided to insure economical use of the line. All these disadvantages are overcome if the power is transmitted in direct-current form. At the receiving end of the transmission

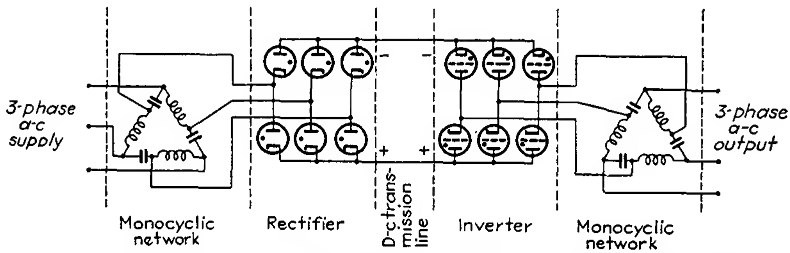


FIG. 173.—Rectifier-inverter combination used for transmitting power by direct current, with alternating-current generation and consumption.

line it is necessary to convert the direct current back into alternating current for local distribution and utilization. The complete process (see Fig. 173) involves generation of high-voltage as alternating current which is then rectified in diode or triode multiphase circuits, smoothed with filter reactors, and sent as direct current along the transmission line. At the receiving end, synchronous gas-triode inverters are used to transform the direct current into alternating current at any desired frequency, in which form it is sent out over the local distribution circuits.

This form of power transmission is in its infancy, there being only one such line in operation in this country, terminated at the General Electric Laboratories at Schenectady. It is believed that the new system will find wide application as soon as its performance under all forms of service has been investigated. One problem is the formation of arcs when insulation fails, or from other causes. An alternating-current arc is easy to extinguish, since the voltage passes through zero twice per cycle, but in

direct-current transmission the arc tends to persist until the voltage is removed. In the Schenectady line a constant-current system is used, so that no more than normal current can flow in the line under any condition, including a direct short circuit, and the arc current is thus limited. The constant-current system has other advantages in transmission, but constant voltage is required in the distribution circuits. The transformation from constant voltage to constant current and back again is made in monocyclic networks, combinations of inductance and capacitance connected as shown in Fig. 173.

Problems

1. Calculate the average power delivered to the load in the full-wave rectifier circuit shown in Fig. 158. The total secondary r-m-s voltage is 650 volts. The rectifier (gaseous type) drop is 15 volts. The load is 1000 ohms.

2. Prove that the no-load voltage provided by a voltage-doubling rectifier (Fig. 160) is $2\sqrt{2}$ times the r-m-s input voltage.

3. The phase difference between the current and the voltage in a series R and L circuit is $\phi = \tan^{-1}(2\pi fL/R)$. Plot the relationship between the value of R and the power output in the circuit shown in Figs. 164 and 167.

4. In the single-tube inverter shown in Fig. 169, the tube fires at 68 volts. The time of discharge thereafter to the extinction voltage (5 volts) is 0.01 sec. Calculate the frequency of operation. Line voltage = 105 volts. R = 100,000 ohms, $C = 0.5\ \mu$f.

Bibliography

HENNEY, KEITH: "Electron Tubes in Industry," McGraw-Hill Book Company, Inc., New York, 1937, Chaps. IV, VII.

MACARTHUR, E. D.: "Electronics and Electron Tubes," John Wiley & Sons, Inc., New York, 1936, Chap. VIII.

DOW, W. G.: "Fundamentals of Engineering Electronics," John Wiley & Sons, Inc., New York, 1937, Chaps. XXI, XXII.

HERSKIND, C. C.: Grid Controlled Rectifiers and Inverters, *Elec. Eng.*, June, 1934.

FOOS, C. B.: Vacuum Tube Controlled Rectifiers, *Elec. Eng.*, April, 1934.

LIVINGSTON and LORD: The Single-tube Thyratron Inverter, *Electronics*, April, 1933.

GULLIKSEN and VEDDER: "Industrial Electronics," John Wiley & Sons, Inc., New York.

REICH, H. T.: The Thyratron Relaxation Oscillator and Its Applications, *Rev. Sci. Instruments*, October, 1933.

WILLIS, BEDFORD, and ELDER: Constant Current D. C. Transmission *Elec. Eng.*, January, 1935.

THOMPKINS, F. N.: The Parallel Type Inverter, *Elec. Eng.*, April, 1933.

CHAPTER XIII

ELECTRONIC COMMUNICATION CIRCUITS

Introduction.—The communication of intelligence and entertainment by electricity involves the transmission of alternating currents which follow the variations in the original intelligence-conveying medium. In the transmission of sound, for example, the alternating current is produced by a microphone which responds to the changes in air pressure. In general the alternating current produced in a communication circuit is of very irregular shape. This irregularity is reduced to order by considering the alternating-current wave, however complex, to be the sum of a number of pure sine-wave components, each having a specified amplitude, frequency, and phase relation to the others. From this point of view, a communication current is the sum of many alternating-current components, and the suitability of a communication circuit to transmit the current is measured in terms of its ability to transmit these alternating-current components simultaneously, without impairing the relative amplitude, the frequency, and the phase relationship of each. The ideal communication circuit would transmit alternating current of any frequency, at any amplitude, and would introduce no relative phase delay. This ideal is, of course, unattainable; the degree to which it is approached in practice depends on the type of communication involved.

61. Amplitude-frequency Response of Communication Equipment.—Since any communications circuit must be capable of transmitting simultaneously many alternating-current components of different frequency, an important characteristic of the circuit is the range of frequencies it can transmit. The frequency range involved becomes wider, the more information or "detail" is transmitted per second. In Table XI are given the frequency ranges required for several important types of service. It will be noticed that even the moderate requirements of simple voice transmission involve a frequency range of 10 to 1, while

in television the range is a million to one. This wide range puts a severe burden on the transmission system. Fortunately in the transmission of voice and music only amplitude and frequency are important; the phase relationship may be disregarded.

TABLE XI.—FREQUENCY RANGES OF VARIOUS COMMUNICATION SERVICES
Audio and Video Services

Type of service	Lower frequency limit, cycles per second	Upper frequency limit, cycles per second
Slow-speed telegraph	0	15
High-speed telegraph	0	80
Wire telephony (speech)	250	2,750
Wire telephony (music)	40	7,000
Sound motion pictures	40	10,000
Limits of audibility	16	20,000
Television (441 lines, 30 pictures per second)	30	2,500,000
Carrier Services		
Carrier telegraph	3,000	10,000
Carrier telephony	4,000	30,000
Radio (long wave)	10,000	300,000
Radio (medium wave)	300,000	1,500,000
Radio (short wave)	1,500,000	30,000,000
Radio (ultrashort wave)	30,000,000	600,000,000

The criterion of a communication circuit, then, is its ability to transmit alternating currents of any frequency within a given range, and to preserve the relative amplitude of each. Note the word *relative* amplitude; any communication circuit must waste power in heating the wires and insulation, and this loss of power means a loss in the over-all amplitude of the wave. But so long as the relative amplitude of each alternating-current component is maintained in relation to the others, the intelligence is preserved.

This requirement of transmitting a wide range of frequencies without discrimination applies to every unit in the communications circuit, including the pickup device or transmitter, the transmission circuit, and the reproduction device or receiver.

It is customary to express the utility of each of these units, and of their component parts, in terms of "amplitude-frequency-response curves." These curves show the frequency range capable of being handled, as well as the amount of amplitude discrimination displayed by the equipment within this range. Typical examples of these response curves are given in Fig. 174. Two of the curves compare the performance of two microphones. Curve A is that of an ordinary telephone transmitter; it will

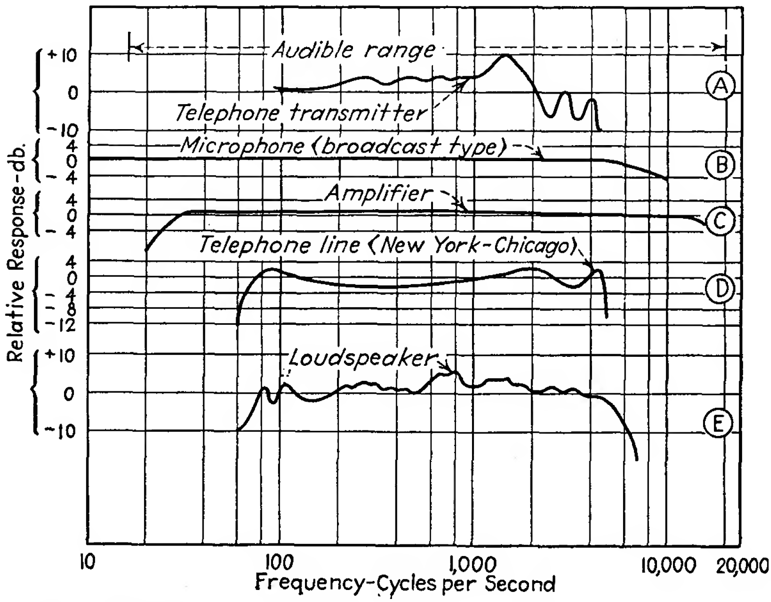


FIG. 174.—Response curves (voltage amplitude in db vs. frequency) of typical communication-circuit elements.

be noted that the low and high frequencies are not transmitted. The voice reproduced at the receiving end thus has weak bass and sibilant sounds. Curve B, of a broadcast-station-type microphone, transmits a much wider range, with less discrimination between frequencies; the voice transmitted in this case is full in the low, medium, and high registers.

The ranges in the scales of these response curves deserve comment. Since the frequency range is very great, a logarithmic scale, which is compressed more and more as the range increases, is used to accommodate the range and to give detail at the

important low frequencies. The vertical scale, the relative response, is usually compressed in the same way, but by the use of a logarithmic unit rather than by the use of logarithmic scale divisions.

The range of power levels encountered in communication circuits is enormous. Even in ordinary speech, the loudest essential speech sounds are at least a thousand times stronger than the weakest essential sounds, and in a symphony orchestra, with its large dynamic range, the ratio may be a million to one. Communication equipment must therefore be able to preserve the relative amplitude of different frequencies, and it must also be capable of responding to changes in over-all amplitude, over wide limits. For convenience in expressing these power-level changes, a logarithmic unit, the *decibel* (abbreviated db), has been devised. The decibel expresses a change in power level as follows:

$$\text{Decibels difference between } P_1 \text{ and } P_2 = 10 \log_{10} \frac{P_1}{P_2} \tag{77}$$

where P_1 and P_2 are two power levels, expressed in the same units. If P_1 is larger than P_2, then the number of decibels is positive, otherwise it is negative.

In communications practice, relative *voltage* levels (in contrast to power levels) are also of interest. Since the power delivered to a given impedance varies with the square of the voltage, the equivalent of Equation (77), expressed in voltage is

$$\text{Number of decibels between } E_1 \text{ and } E_2 = 10 \log_{10} \frac{E_1^2}{E_2^2} = 20 \log_{10} \frac{E_1}{E_2} \tag{78}$$

This equation represents a change in *power level* only when the two impedances (across which the voltages develop) are identical in amplitude and phase. However, Equation (78) may be applied simply to express difference in voltage level, without regard to power. This is particularly convenient in connection with voltage amplifiers in which the input power is very small and usually of no consequence. When used to express voltage changes only, Equation (78) should read "decibels voltage

gain," not "decibels." This distinction should always be stated.

62. Single-stage Audio-frequency Amplifiers.—The range of frequencies from 10 to 20,000 cycles per second is called the audio-frequency range, since these are, approximately, the limits of audible vibration. Amplifiers suitable for transmitting all or part of this range are called audio-frequency amplifiers. The basic amplifier connection, Fig. 175, is fundamental to all such amplifiers. It contains a triode, tetrode, or pentode tube, a source of direct grid-bias voltage (usually a grid-bias filter), a grid impedance across which the signal voltage is developed, a source of direct plate-bias voltage, and a plate impedance across which the amplified signal voltage is developed.

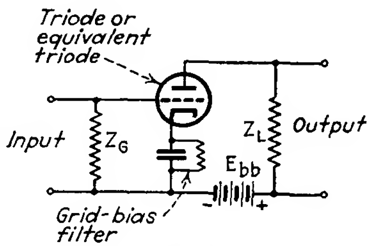


FIG. 175.—Basic amplifier connection, employing bias filter and applicable to triodes, tetrodes and pentodes.

The frequency response of the amplifier depends primarily on the nature of the grid and plate impedances. If these parameters are simple resistances, then the impedance displayed by them to the signal currents (see page 231, Sec. 52) is the same for all frequencies and there is no frequency discrimination. On the other hand, if the grid impedance is a pure capacitance, it possesses low impedance for high frequencies (page 233), and the low frequencies are thereby emphasized relative to the high. Likewise, if the grid or plate impedances are inductances, the impedance is high for high frequencies (page 232), and the high frequencies are emphasized. To avoid frequency discrimination, resistance is almost universally used as the grid impedance, and the plate impedance is either a resistance or a high inductance and resistance in series.

Distortion.—An important consideration in amplifiers is the distortion they introduce to the signal. Whenever the wave form of the input signal is not reproduced faithfully in the output, the change in wave form is called *distortion*. Actually this distortion amounts to the insertion by the amplifier of spurious frequency components. When a pure sine wave is transmitted through an amplifier, these spurious components

are always whole multiples (harmonics) of the basic sine-wave frequency and hence are called harmonic-distortion components.

Distortion is associated with the dynamic characteristic of the tube and circuit, as shown in Fig. 176. If the dynamic characteristic is a straight line, then the output current is always proportional to the input voltage and no change in wave form can occur within the tube. To insure a straight dynamic characteristic, it is important that the tube be operated so that its bias point (page 247) is in the midst of the "linear" region

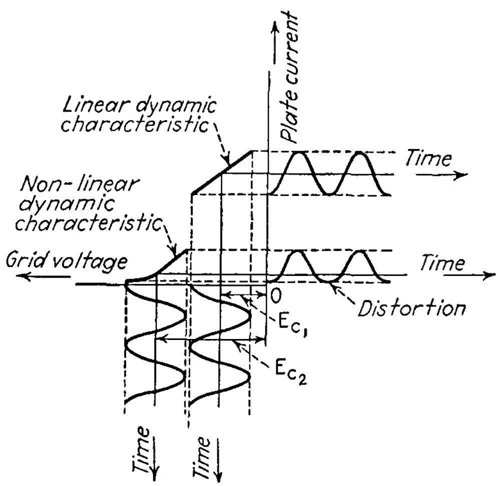


FIG. 176.—Distortion of wave form caused by curvature of the dynamic characteristic. The grid-bias voltage E_{c_2} is excessive, whereas E_{c_1} permits operation in the linear region of the characteristics.

of the static tube characteristics, *i.e.*, where the curves are approximately straight lines. Since the position of the bias point is determined by the applied grid and plate voltages, these voltage values must be properly chosen if distortion is to be minimized. Furthermore, the voltage range of the input voltage (grid-voltage "swing") must not be so great that the excursions on either side of the bias point run into the nonlinear regions of the static curves, since this will produce a bend in the dynamic characteristic, with resulting distortion of wave form. To avoid these wide excursions the maximum amplitude of the grid-voltage signal must always be less than the value of the direct grid-bias voltage.

Such an amplifier, whose grid-bias voltage lies in the region of the linear portion of the static characteristics, and whose grid-signal swing is not allowed to exceed the bias value (*i.e.*, whose grid never becomes positive) is called a class *A* amplifier. It is universally used in single-tube audio-amplification practice. It has characteristics of low distortion and low efficiency.

Effect of Tube Capacitances.—A limiting factor in the frequency response of an audio amplifier is the capacitance displayed by the tube elements themselves. The grid-cathode capacitance shunts the grid impedance and lowers its response to high frequencies, and the same effect is true of the shunting effect of the plate-cathode capacitance on the plate impedance. The capacitance between grid and plate acts to transfer energy from the plate to the grid, and acts as the self-excitation agency in certain oscillator circuits. In audio-amplifier circuits the tendency to oscillate is not great since the grid and plate impedances are not frequency-selective. When high amplification is desired, screen-grid or pentode tubes are used since their grid-plate capacitance is low.

Voltage Versus Power Amplifiers.—The single-tube amplifier has thus far been treated as a voltage amplifier, *i.e.*, one which accepts a weak signal voltage from the grid impedance and develops an amplified signal voltage across the load impedance. If the load impedance is in itself a power-utilizing device (such as a loud-speaker or other reproducing unit), the important quantity is not voltage but power, the product of voltage and current. In such cases the amplifier is called a power amplifier, and the tube used must be capable of delivering large values of plate current and of accommodating large variations in this current. Furthermore, the efficiency of this type of amplifier is important since the power delivered by a tube of given size depends on efficiency. The class *A* amplifier, being inefficient, is not particularly well suited from this point of view, yet it is the only type of single-tube amplifier with low-distortion characteristics. The alternative is a double-tube or "push-pull" amplifier, a single stage of amplification containing two tubes, which work on alternate half cycles of the signal. This mode of operation permits higher efficiency.

Class B Amplification.—The type of amplification employed in the push-pull circuit depends on the grid-bias-voltage values

used on each tube. If the bias value corresponds to the linear portion of the static characteristic curves, then the amplifier operates in class *A* and no increase in efficiency is gained. But if the grid bias is such that in the absence of any signal, no plate current flows at all (the grid bias is at its plate-current cutoff

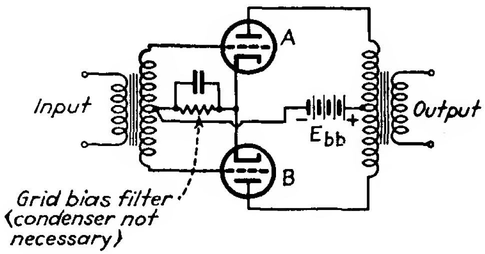


FIG. 177.—Push-pull amplifier connection.

value), then the amplifier operates in class *B*. This mode of operation manifestly reduces the average plate current relative to the amplitude of plate-current swing, and thus increases the efficiency, but at the same time it produces serious distortion,

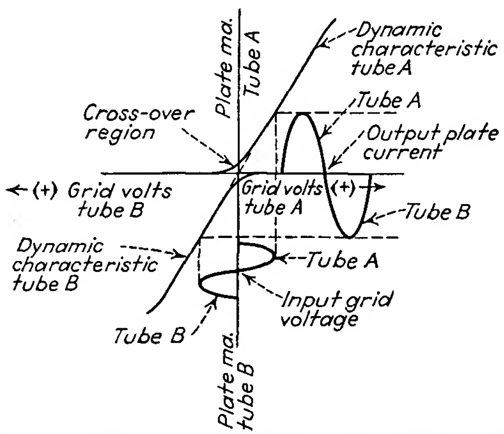


FIG. 178.—Dynamic operation of a class *B* push-pull amplifier stage (*cf.* Fig. 177).

at least in a single-tube amplifier. Under this condition all the negative half cycles of the grid voltage produce no change in plate current at all, and the output current has a badly distorted shape. But if two tubes are used in push-pull so that one tube supplies the half cycles missing from the plate current of

the other (see Figs. 177 and 178), then the net output current (the sum of the two plate currents) has the same form as the input signal and the distortion is nullified. In this manner the advantage of higher efficiency of operation is gained without introducing unwarranted distortion. The distortion is somewhat higher than in the class *A* type but in general the distortion is tolerated in view of the high efficiency.

To take full advantage of the high-efficiency characteristic of the class *B* amplifier it is usual to apply a strong grid-signal voltage, thus making the plate current variation (signal) large compared with the average plate current. This large grid signal is sufficient to overcome the grid bias (large as the bias is) and "drive" the grid positive. Grid current flows as a result, and the power consumed by the grid input circuit may be considerable. If the signal source (a preceding amplifier stage in the usual case) can supply this power, no serious distortion results from the grid-current flow.

The class *B* amplifier, as applied to audio frequencies, is always a double-tube push-pull arrangement, each tube with grid biased to plate current cutoff, and operated with a grid signal whose maximum amplitude exceeds the grid-bias voltage. Its characteristics are higher efficiency and somewhat greater distortion than the corresponding class *A* amplifier.

63. Coupled Audio Amplifiers.—The single-stage amplifiers thus far considered are limited in the voltage and power amplification they can perform. If Equations (44) and (71) are applied to the tubes available for audio amplifiers, it will be discovered that the maximum voltage gain is about 25 (28 db of voltage gain) for a triode and 1000 (60 db of voltage gain) for a pentode, and that the maximum power is about 1 watt per r-m-s volt on the grid (a pentode in class *B*). To obtain large power output from a small voltage source, it is necessary to employ several stages of amplification one after the other, the output of one feeding the input of the next. In this way total voltage gains of over 150 db of voltage can be obtained, and over-all power level increases of 70 db (10 million times) realized. In coupled amplifiers it is necessary, of course, to consider the frequency response of the coupling connection, since any frequency discrimination in this part of the amplifier will be carried thorough all succeeding stages to the final output.

Three important methods are used for coupling audio-frequency amplifiers: (1) Resistance-capacitance coupling; (2) impedance-capacitance coupling; (3) transformer coupling. A fourth method, direct coupling, is of no practical importance in communication practice, but it does serve to introduce the logic of coupled amplifiers, so it will be described.

The direct-coupled amplifier obtains the grid-signal voltage for one tube directly from the plate impedance of the preceding tube, as shown in Fig. 179. Consequently, if the input voltage is amplified faithfully by the first stage, it is passed on unimpaired to the second stage, where it is amplified further. There is, therefore, no distortion or frequency discrimination exercised other than that in the individual stages themselves. Consequently the direct-coupled amplifier is a very high-grade amplifier. But it suffers from one important drawback, the necessity of a plate-voltage supply equal to the sum of all the plate voltages required for all the stages. This requirement is equivalent to having a separate anode supply for each stage. If a common plate-voltage supply is used for all tubes, the direct connection between stages applies the plate voltage of one tube to the grid of the following tube. For extremely low-frequency amplification (below 30 c.p.s.), however, other methods of coupling fail, and the direct method must be used with a separate plate-voltage supply for each tube. The need for such low-frequency amplification in communication practice is fortunately not great.

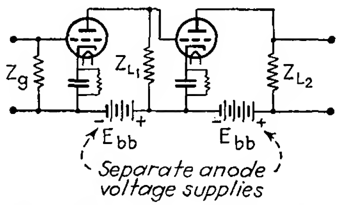


FIG. 179.—Direct-coupled two-stage amplifier.

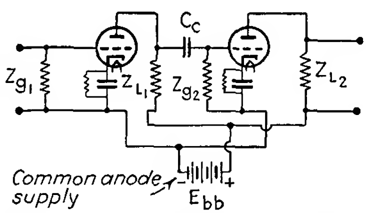


FIG. 180.—Resistance-capacitance-coupled amplifier.

The Resistance-capacitance-coupled Amplifier.—The resistance-capacitance-coupled amplifier, shown in Fig. 180, is a modification of the direct-coupled amplifier, in which a coupling capacitor C_c is used to isolate the plate of one tube from the grid of the following tube, so far as direct current is concerned. The direct plate voltage of the first stage does not affect the grid

bias of the second, and hence the same plate-voltage supply may be used for both tubes. The coupling capacitor acts to pass the alternating voltage drop from the plate load resistor to the grid resistor of the following tube, but since the impedance of the capacitor decreases as the frequency increases, the high-frequency alternating-current components are transferred with less loss than the low-frequency components. In the resistance-capacitance-coupled amplifier, therefore, the low-frequency response is limited by the size of the coupling capacitance. In practice values from 0.01 to 0.1 μf give good response down to 50 c.p.s. If amplification at lower frequencies is desired, a larger capacitor may be used, but the possibility then remains that the capacitor charge accumulated from a large signal will not leak off rapidly enough through the grid- and plate-resistance path. A direct voltage is thereby temporarily trapped across the coupling capacitor, and this direct voltage biases the grid of the following tube to plate-current cutoff, thereby "blocking" it. After a short time the capacitor charge leaks off, to be regained on the next large signal pulse. The circuit thereby starts and stops periodically, an effect usually called "motor boating," from the "put-put" noise produced. The principle of relaxing a condenser charge periodically is used in the frequency-division circuits described in Sec. 69.

The complete analysis of the resistance-capacitance-coupled amplifier at low, medium, and high frequencies is complicated, but it can be handled fairly simply if it is assumed that the coupling capacitor has no impedance for the medium and high frequencies, and that the shunt capacitances (from the tube elements and the wiring) across the grid and plate impedances have no effect on the low frequencies. An equivalent impedance $R_{\text{eq.}}$ is formed equal to the dynamic plate resistance, the plate-load resistance, and the grid resistance of the following tube, all in parallel. The equations for the gain per stage (gain $= e_{g2}/e_{g1}$) are:

At low frequencies:

$$\text{Gain} = \frac{g_m R_{\text{eq.}}}{\sqrt{1 + (X_c/R)^2}}. \tag{79}$$

At mid-range frequencies:

$$\text{Gain} = g_m R_{\text{eq.}}. \tag{80}$$

At high frequencies:

$$\text{Gain} = \frac{g_m R_{\text{eq.}}}{\sqrt{1 + (R_{\text{eq.}}/X_s)^2}}, \tag{81}$$

where

$$\frac{1}{R_{\text{eq.}}} = \frac{1}{r_p} + \frac{1}{R_p} + \frac{1}{R_g}. \tag{82}$$

$$R = R_g + \frac{r_p R_p}{r_p + R_p}. \tag{83}$$

$$X_c = \frac{1}{2\pi f C_c}. \tag{84}$$

$$X_s = \frac{1}{2\pi f C_s}. \tag{85}$$

and where C_s is the total shunting capacitance, equal to the plate-

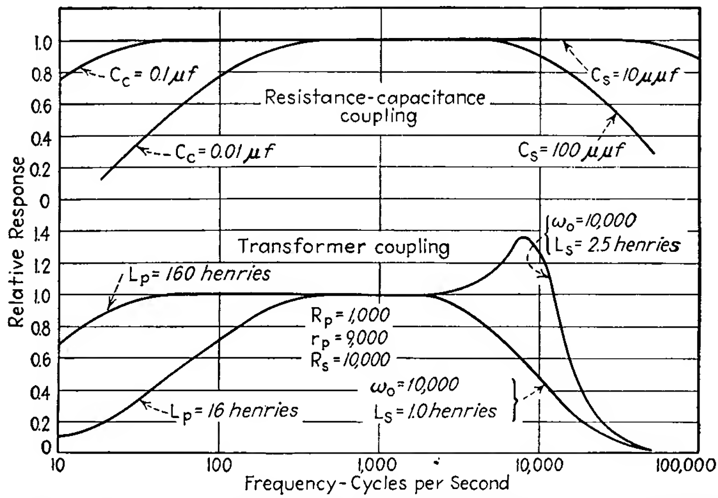


FIG. 181.—Response curves of resistance-capacitance-coupled and transformer-coupled amplifiers, computed from Eqs. (79) to (81), and (86) to (88).

cathode capacitance of the first tube, the grid-cathode capacitance of the second tube, plus wiring capacitance.

Typical response curves calculated according to these formulas are given in Fig. 181. It will be noted that for good high-frequency response, the shunting capacitance must be kept at a minimum.

The Impedance-capacitance-coupled Amplifier.—The disadvantage of using a resistance as a load impedance is the fact that a

large direct-voltage drop is developed across it by the flow of the static direct plate current, and that this voltage drop is subtracted from the voltage of the plate-voltage supply, the remainder being supplied to the tube itself. If too large a plate-load resistor is used, the voltage drop across it becomes excessive and the plate current falls to a low value. Yet high plate impedance is required to obtain a reasonable gain from the stage.

A compromise between these two effects is secured by the use of a high inductance as the plate-load impedance. The resistance of this inductance may be made small, so that the direct-voltage drop across it is only a few volts. Its impedance to alternating current is high, however, and it increases as the frequency of the alternating current increases. This frequency discrimination can be tolerated, if the value of the inductance is large enough to develop sufficient voltage at the lowest desired frequency, since the high-frequency response is lowered by the effect of the shunting capacitance in the circuit.

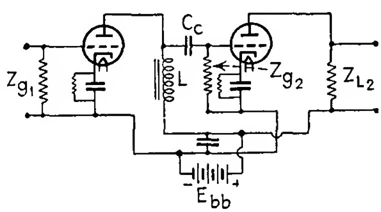


FIG. 182.—Reactance-capacitance-coupled amplifier.

The gain of an impedance-capacitance-coupled amplifier may be calculated with sufficient accuracy for ordinary purposes by applying Equations (79), (80), and (81), but the value of $R_{eq.}$ must be replaced by $Z_{eq.}$ which is the parallel impedance of r_p, X_L, and R_g, and R must be replaced by Z, which is the sum of the grid impedance in series with the X_L and r_p values in parallel. Impedance amplifiers are not so widely used as the resistance-coupled and transformer-coupled types.

Transformer-coupled Amplifiers.—A very widely used method of coupling audio-frequency amplifiers is transformer coupling. The transformer has two windings which are isolated so far as direct current is concerned, and hence no coupling capacitance is required. Furthermore, the primary winding of the transformer, used as the plate-load impedance, has low resistance and high reactance, and thus possesses the advantages of the impedance load as against the resistance load outlined above. Finally, the design of the coupling transformer contains many variables which can be adjusted to give almost any desired type

of frequency response up to 200,000 cycles per second and down to 30 cycles per second. The transformer may be designed to possess a certain amount of voltage amplification within itself, by incorporating more turns on the secondary winding than on the primary. The voltage step-up ratio is of use in obtaining a high over-all gain in a multistage amplifier.

The amplification per stage of a transformer-coupled amplifier depends to some extent on the transformer design. In the middle range of frequencies the principal determining factors are the step-up ratio n, and the amplification factor μ of the tube. The mid-range gain is then:

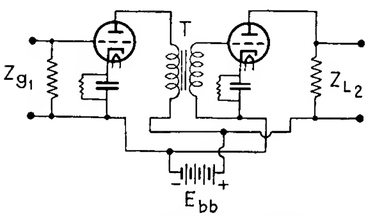


FIG. 183.—Transformer-coupled amplifier.

$$\text{Gain} = \mu n \tag{86}$$

In the low-frequency range, the principal limitation is the inductance L_p of the transformer primary. The low-frequency equation is:

$$\text{Gain} = \frac{\mu n}{\sqrt{1 + (X_L/R_{pt})^2}}, \tag{87}$$

where $X_L = 2\pi f L_p$, and R_{pt} = dynamic plate resistance plus transformer-primary resistance.

At the high-frequency end of the spectrum, the amplification is affected by a tendency to resonate between the leakage inductance L_s of the transformer and the shunting capacitance C_s of the tubes and wiring. The resonant frequency of this combination by Equation (68) is

$$f_0 = \frac{1}{2\pi\sqrt{L_s C_s}}$$

The high-frequency amplification is given by:

$$\text{Gain} = \frac{\mu n}{\sqrt{\frac{f(R_{pt} + R_s)^2}{f_0(2\pi f_0 L_s)^2} + \left(\frac{f^2}{f_0^2} - 1\right)^2}}, \tag{88}$$

where R_s is the secondary resistance divided by the square of the step-up ratio. This equation is rather complicated, but

generalized curves can be prepared from it which make computation fairly simple. Typical examples are given in Fig. 181.

The transformer parameters (primary and secondary resistances, primary inductance, and leakage inductance) are seldom stated by the manufacturer of the transformer, but they can be measured by a straightforward method. In practice, the transformer step-up ratio and its frequency-response curve are specified under given operating conditions (particularly the input-tube resistance, the direct-current flow in the primary, and the value of the grid resistor, if any, shunted across the secondary); this information serves to give the performance of the stage without any involved calculations.

In general, transformer-coupling is used with tubes having direct plate currents larger than 10 ma.; resistance coupling with tubes having less than this value, although exceptions to this rule are frequently found.

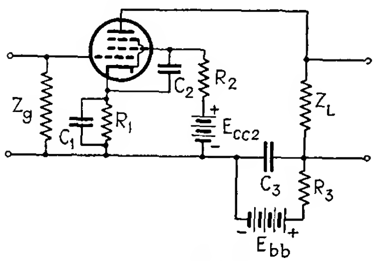


FIG. 184.—By-passing filters in a pentode-amplifier stage.

By-passing Practice in Triode, Tetrode, and Pentode Amplifiers. In all amplifiers it is desirable to develop the maximum alternating signal voltage across the plate-load impedance. This not only insures maximum gain for the stage, but it also prevents the amplified signal from being applied to other parts of the circuit, notably the grid circuit of the tube, where it would do harm. In the discussion of the grid-bias filter (Sec. 53), it was pointed out that the filter capacitor serves to prevent a signal voltage from developing across the bias resistor, and thus causes the full signal voltage to appear across the plate-load impedance. However, there are other parts of the circuit where similar treatment is needed. The plate-voltage supply, for example, usually possesses appreciable internal resistance, across which appears a signal voltage in response to the changes in plate current. If this internal resistance is not part of the plate-load impedance, then the signal developed across it is lost, and more important, this voltage will be applied to the plate circuits of all other tubes connected to the plate-voltage supply. This is a very potent source of trouble in amplifiers, since it

leads to the excitation of oscillations in the circuits. Combinations of resistance and capacitance, called decoupling filters (see Fig. 184), are commonly employed in each stage to prevent this effect.

In tetrode and pentode tubes still another possibility of improper signal voltage is presented by the screen-grid circuit of the tube. The proper function of the screen grid is obtained only when its potential with respect to cathode remains constant. Yet the screen grid can readily act as a "plate," and the screen current will develop signal variations if means are not provided to prevent it. Usually a capacitance is connected directly between the screen grid and the cathode to bring them to the same potential so far as the alternating-current signal components are concerned, while still allowing the direct-current difference of voltage on which the screen-grid action depends. The action of the capacitor, commonly called "by-passing," is illustrated in Fig. 184.

64. Principles of Carrier Communication.—The audio-frequency equipment described in the preceding section suffers from two basic limitations. In the first place, any two-wire audio-frequency communication circuit is limited to one message at a time. If two or more messages are sent over the circuit at once, their mutual interference renders them unintelligible. In the second place audio-frequency signals cannot be sent directly into space, in the form of radio waves, since the radiation of energy from an antenna takes place economically only at frequencies higher than 30,000 c.p.s., which is well out of the audio-frequency range. Both of these limitations are removed by the use of *carrier communication.* Carrier communication involves the use of a high-frequency alternating current (or "carrier") as the basic communication agent, and the modification of this high-frequency current by comparatively low-frequency currents which contain the intelligence to be transmitted.

Modulation, the Modification of the Carrier Frequency.—Since the carrier current is a sine-wave alternating current, it possesses three aspects which may be varied periodically: its amplitude, phase, and frequency. All three aspects have been used in practical carrier transmission systems, but only amplitude is widely used at present. The modification of a high-frequency current by a low-frequency one is called modulation; when

the amplitude is modified, the process is called *amplitude modulation.*

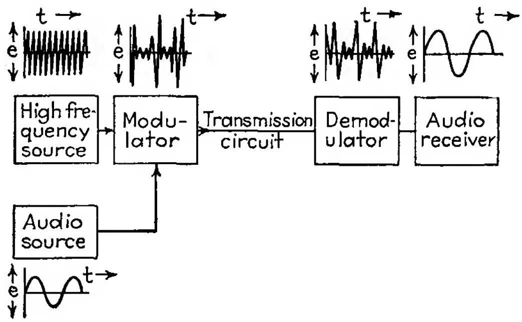


FIG. 185.—Elements of a carrier-communication system.

To illustrate the process of amplitude modulation consider a sine-wave voltage whose frequency is, say, 1,000,000 c.p.s.

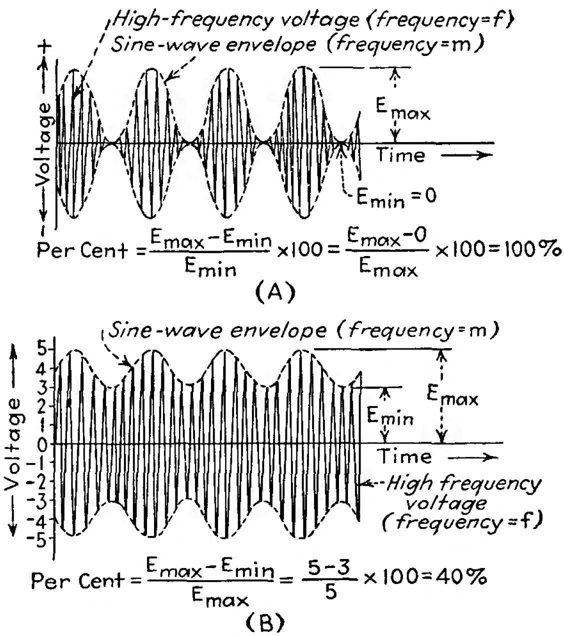


FIG. 186.—Amplitude modulation.

Now, if the power source which creates the wave is caused to change its power output (but not its frequency) at an audio-frequency rate of, say, 1000 c.p.s., the amplitude of the sine-wave

voltage will change in accordance with these power changes. If the audio-frequency signal (called the modulating signal) has a sine-wave form, then the changes in amplitude undergone by the carrier wave will have a sine-wave form. A typical wave form so produced is shown in Fig. 186. In the example cited, there are 1000 cycles of the carrier for every cycle of the modulating signal. If an imaginary line be drawn through the maximum height of each carrier cycle, this line (called the envelope of the carrier) will have the shape of the modulating signal.

If the changes in amplitude are as great as possible, as shown in Fig. 186*A*, the carrier is said to be 100 per cent amplitude-modulated, whereas if the changes in amplitude are less (corresponding to less violent changes in the strength of the power source) the amplitude modulation is less than 100 per cent. The percentage modulation is given by the formula

$$\text{Percentage modulation} = \frac{E_{\text{max.}} - E_{\text{min.}}}{E_{\text{max.}}} \times 100 \text{ per cent} \quad (89)$$

It will be seen that the modulated-carrier system includes two frequency components, that of the carrier and that of the modulating signal, the first carrying the intelligence of the second. Let the carrier frequency be f c.p.s. and the modulating signal frequency m c.p.s., and let the percentage modulation be $k \times 100$ per cent. The equation for the resulting modulated carrier voltage, e, is

$$e = E_{\text{max.}}(1 + k \cos 2\pi mt)(\cos 2\pi ft). \quad (90)$$

The same form of equation applies, of course, if a modulated current is considered rather than a modulated voltage. The phase difference between f and m has been neglected for simplicity.

If the methods of trigonometry are applied to multiply the two cosine terms given above, a rather curious result is found. The voltage e is found to be the sum of three voltage, one of carrier frequency f, one having a frequency $f + m$, and another having a frequency $f - m$. The transmission of an amplitude-modulated carrier thus involves three high-frequency currents, the carrier and two "side-band" frequencies which are separated from the carrier by a frequency interval equal to the modulating frequency. If the modulating signal consists of several simul-

taneous sine-wave components, each signal component gives rise to two individual side-band frequencies which cluster about the central carrier frequency.

This side-band analysis reveals another important aspect of carrier transmission, namely, that the transmission circuit must respond to currents of several different frequencies, including the carrier frequency and several frequencies slightly higher and lower than the carrier. In fact, the range of frequencies involved (the difference between the highest and lowest side-band frequency) is twice as great as the corresponding range of frequencies in the original audio-frequency-modulating signal. But the *percentage* range of frequencies is very small in the carrier case compared with the original audio signal. Good-quality voice transmitted directly by audio-frequency methods must have a frequency range from 60 to 6000 c.p.s., a range of 100 to 1. If impressed on a carrier frequency of 1,000,000 cycles per second (equal to the central frequency of the broadcast-station region) the same signal would entail frequencies from 994,000 to 1,006,000, a ratio of only 1.012 to 1. Consequently, in carrier transmission highly frequency-responsive circuits may be used. Furthermore, by selecting different carrier frequencies at intervals and modulating each with a different signal, several messages may be sent simultaneously over the same circuit. The only requirement is that the spacing between the separate carrier frequencies be equal to the total side-band range involved in each case, so that the side bands of one carrier do not interfere with the side bands of an adjacent carrier. This is the principle by which several messages are sent simultaneously over a single pair of telephone wires, and by which many different radio stations may use the ether simultaneously without interference.

Demodulation: Separating the Signal from the Carrier.—When a carrier current arrives at the receiving end of the communication circuit, two important steps must be taken to restore the original intelligence in a form to which the ear will react. The first of these steps is *tuning*, the separation of the desired carrier frequency from other carrier frequencies which may be present in the circuit, a process carried out by means of a tuned circuit whose resonant frequency is equal to the desired carrier frequency. The second process is *demodulation* by which the audio-frequency signal is separated from the carrier frequency.

Demodulation, more accurately, is the translation of the carrier frequency and side-bands frequencies into the original audio-frequency alternating-current components, corresponding to the signal employed to modulate the carrier at the transmitter.

The process of demodulation may be explained graphically or by means of trigonometry. The graphical method is illustrated in Fig. 187. The modulated carrier voltage is applied to a device which passes current in one direction more easily than in the other, *i.e.*, to some type of rectifier. The characteristic of an ideal rectifier is shown. The output current resulting from the application of this voltage has exactly the

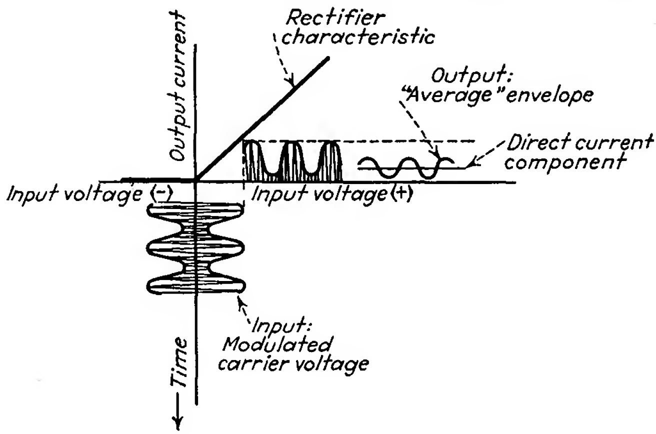


FIG. 187.—Demodulation of an amplitude-modulated carrier by an ideal rectifier.

same form as the voltage, except that the negative half cycles are missing. The result is a current which has a varying direct-current component in addition to several alternating-current components. If this rectified current is passed through a device possessing electrical inertia (an inductance, for example), the device will not respond to the individual half cycles of rectified carrier current, but will respond to the average level of these half cycles, *i.e.*, the varying direct-current component whose average level changes as the amplitude of the modulated carrier changes. Consequently the current in the receiving device has the form shown in Fig. 187. It is a reproduction of the envelope of the carrier, which, in turn, is a reproduction of the original audio-frequency signal at the transmitter. If a condenser is shunted across the receiving device, the high-

frequency currents (the "remnants" of the carrier) are prevented from actuating the receiver, by the by-passing action referred to in Sec. 63. The process thus has the over-all affect of removing the audio-frequency components from the carrier, and then removing the carrier, leaving only the audio-frequency signal, which thereupon actuates the reproducing device (loud-speaker or telephone receiver), after amplification to the required power level.

Mathematically, the process of demodulation is best shown by the use of the equation of the square-law detector (demodulator). This device is commonly used in radio practice; essentially it is a diode or triode tube acting with a load resistor, so that its output current increases as the square of the input voltage. Consequently the positive peaks of the modulated carrier are emphasized relative to the negative peaks, and the action of demodulation is performed. The equation of the square-law detector is

$$i_{\text{output}} = De_{\text{input}}^{2} \tag{91}$$

If the input voltage is the modulated-carrier voltage represented by Equation (90)

$$e_{\text{input}} = E_{\text{max.}}(1 + k \cos 2\pi mt)(\cos 2\pi ft), \tag{90a}$$

then the output current is

$$i_{\text{output}} = DE_{\text{max.}}^{2}[(1 + k \cos 2\pi mt)(\cos 2\pi ft)]^2. \tag{92}$$

The expansion of the last factor in Equation (92) produces terms corresponding to direct current, to the modulating frequency, to twice the modulating frequency, and terms of high frequency (carrier, side bands, and multiples of both). Of these only the modulating frequency and its harmonics are of importance in the receiving device, since the direct current is not a signal, and the high-frequency terms are by-passed. The modulating frequency is, of course, the desired audio-frequency signal, whereas the harmonics constitute distortion. The magnitude of this distortion increases with the percentage modulation, but it is not ordinarily serious because of the tolerance of the ear.

Summary: The Essential Elements of a Carrier Communication System.—The foregoing sections show the necessity of the

following elements in a carrier communication system: (1) a source of high-frequency alternating current, *i.e.*, some form of oscillator; (2) a modulator for varying the amplitude of the carrier current in accordance with the audio-frequency signal; (3) the transmission circuit; (4) a demodulator or detector for separating the audio-frequency signal from the carrier; and (5) amplifiers for increasing the voltage or power level at any point in the circuit, and for isolating various parts of the system from other parts. These five elements include the basic func-

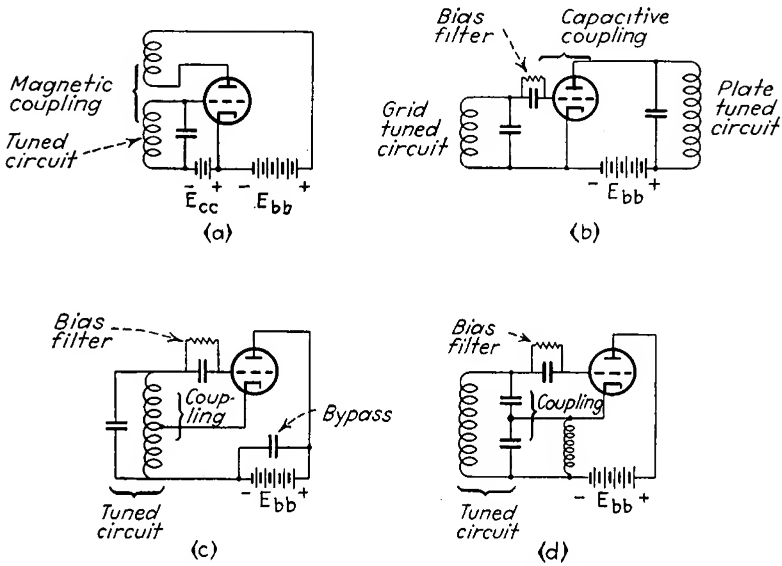


FIG. 188.—Oscillator circuits: (*a*) tuned grid, (*b*) tuned plate, tuned grid, (*c*) Hartley, and (*d*) Colpitts.

tions of electronic tubes in all communication practice: oscillation, modulation, demodulation (including rectification), amplification, and frequency conversion. The circuits used in performing these functions are described in the following sections.

65. Oscillator Circuits.—An oscillator circuit is essentially an amplifier having a parallel-tuned circuit for grid or plate impedance, or both, and having some means of feeding energy from the plate circuit to the grid circuit. The feedback connection can be made in various ways, each giving rise to a particular variety of oscillator circuit. Perhaps the most familiar is the "tickler" oscillator (Fig. 188*a*), having a tuned circuit in the grid and a

simple inductance coil in the plate, the plate coil (tickler) being coupled magnetically to the grid coil. When the plate voltage is applied to the tube, the initial surge of current in the plate coil induces a current in the grid coil, and oscillations are set up in the grid circuit. The alternating voltage thereby applied to the grid is amplified by the tube, and the amplified signal, passing through the plate coil, induces still further oscillations in the grid circuit. The amplitude of oscillation thereby builds up until the equilibrium is reached between the losses in the circuit and the ability of the tube and battery to supply the required alternating power. Use may be made of the oscillating current in the plate coil by coupling a load to it.

Another arrangement employs a tuned circuit in the plate, with inductive coupling to the grid. The frequency of oscillation in this case is given by

$$f_r = f_0\sqrt{1 + \frac{R}{r_p}}, \tag{93}$$

where f_0 is determined by the L and C of the tuned circuit, *i.e.*,

$$f_0 = \frac{1}{2\pi\sqrt{LC}}$$

The correction factor is small, provided that r_p the dynamic plate resistance is large compared with R, the resistance in the plate tuned circuit, as it usually is.

Still another type of oscillator employs tuned circuits in both grid and plate circuits. The tuned-grid, tuned-plate circuit usually relies for feedback on the grid-plate capacitance within the tube, rather than on inductive or capacitive coupling between the coils.

In practical oscillator circuits, one parallel-tuned circuit is often made to serve in both the plate and grid circuits. Such are the Hartley and the Colpitts oscillators, shown in Figs. 188*c* and 188*d*. In each oscillator circuit there is a definite relationship between the g_m and r_p values of the tube used and the L, C, R, and M values employed in the tuned circuits and in the coupling between them. These relationships must be satisfied before oscillation starts.

Frequency Stability of Oscillators.—An important consideration in the use of oscillators as alternating-current gen-

erators for carrier service is their frequency stability. If the frequency tends to vary or "drift" during operation, then the carrier and side bands of one communication channel will begin to interfere with the adjacent channel. The most widely used method of stabilizing frequency is the use of a piezoelectric quartz crystal in the grid circuit of the oscillator tube. This crystal is mechanically ground with respect to its optic axes so that it exhibits a tendency to vibrate mechanically at a very rapid rate when stimulated by an alternating voltage applied across its faces. The quartz generates alternating current in this vibrating process and the alternating current so generated tends to stabilize the grid circuit, that is, the vibration rate of the crystal acts as the basic frequency source of the oscillator. To stabilize the crystal frequency still further, it is common practice to maintain the temperature of the crystal within narrow limits. By these means frequency stability of a few parts in a million is readily obtainable. All broadcast stations, for example, are required by law to maintain their carrier frequencies within 50 cycles of the assigned value, representing an accuracy of about 50 parts in a million. Actually the accuracy maintained is about 5 parts in a million. The outstanding example of accuracy, however, is the quartz-crystal clock which is so accurate (about 1 part in 5 million) that its output can be used as a standard of time having greater precision (over short periods) than any mechanical clock yet built.

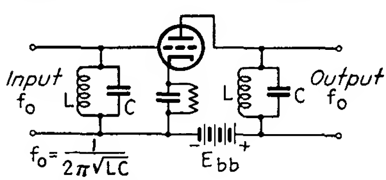


FIG. 189.—Tuned amplifier.

66. Tuned Amplifiers.—If for any purpose a narrow range of frequencies, such as a carrier and its side bands, are to be amplified, the use of tuned grid and plate circuits greatly increases the gain obtained in the amplifier, owing to the frequency-selective properties of these circuits. Consequently, if the output of an oscillator is to be amplified, or if the carrier frequency after modulation is to be amplified, tuned amplifiers are employed. Figure 189 shows a typical example. The resonant frequency of the input and output tuned circuits corresponds to the carrier frequency to be amplified. Steps are taken, usually by the use of a screen-grid tube, to prevent feedback from plate to grid, so that the

amplifier amplifies but does not oscillate of its own accord. The performance of the tuned amplifier is treated by the same methods employed in the audio-frequency amplifier, with important differences due to the narrow range of frequencies involved.

In applying tuned amplifiers to the amplification of a modulated carrier, in which a range of frequencies is involved, it is important that the tuned circuit employed shall pass all the side-band frequencies equally well. Ordinary tuned circuits are apt to be sharply responsive to a narrower range than is common in voice modulation. If two tuned circuits are employed, however, with their coils linked transformer fashion, the reaction of one circuit upon the other broadens the response of the

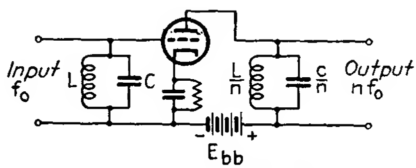


FIG. 190.—Frequency-multiplying tuned amplifier.

combination over a wide band, sufficient to pass all the desired side-band frequencies without discrimination. This type of coupling is commonly called "band-pass" coupling. For extremely wide ranges of side-band frequencies, such as are encountered in television, tuned circuits can be used only if they are specially designed for the purpose.

Frequency-multiplying Tuned Amplifiers.—If the resonant frequency of the plate circuit in a tuned amplifier is made some multiple of the resonant frequency of the grid circuit (Fig. 190), then by the process explained on page 268 the plate circuit will be energized at its own resonant frequency and the frequency of the current "passing through" the amplifier will be in effect multiplied. Usually the multiplication so obtained is limited to two or three times. If further multiplication is required, two multiplying stages may be used, but only if an isolating amplifier is used between them. This isolating or "buffer" stage serves simply to repeat the signal without multiplying it and prevents the signal of the second multiplying stage from acting on the

grid or plate of the first multiplying stage. Frequency division by this means is not possible, for the reasons outlined on page 268.

Classes A, B, and C Applied to Tuned Amplifiers.—Tuned amplifiers may be operated in class A, B, or C, depending on the grid-bias voltages and the excursion of the grid-signal voltage. Ordinarily in radio-receiver practice the efficiency of amplification is not important, so class A is used. But for transmission purposes, especially in high-power stages, efficiency is an important consideration, so classes B and C are used. Class B operation in tuned amplifiers is similar to that in audio amplifiers, the grid being operated with cutoff bias and the maximum amplitude of the grid signal exceeding the bias value. One important difference between radio-frequency and audio-frequency practice appears. The class B audio amplifier must be push-pull in form to avoid excession distortion. In the tuned radio-frequency amplifier, however, a single tube may be used in class B because no appreciable distortion appears across the load impedance. The distortion produced in a sine-wave signal, it will be remembered, consists of components whose frequencies are whole multiples of the signal frequency. The tuned circuits employed in radio-frequency amplifiers have no appreciable impedance to these high-frequency distortion components, and consequently they do not appear across the plate-load tuned circuit or the grid tuned circuit of the following stage. The same reasoning applies to a modulated radio-frequency signal, since the distortion components of each side-band frequency are well out of the response range of the tuned circuits used. Consequently a single tube class B tuned-amplifier stage may be used to amplify a carrier, modulated or unmodulated, without distortion appearing in the output. Such amplifiers are used after the modulator stage in radiotelephone transmitters.

This tolerance of tuned amplifiers to distortion has led to still another class of amplifier, class C, in which high efficiency and high power output are obtained without reference to the distortion produced. The grid in a class C amplifier is biased to a point considerably beyond the cutoff, and the grid signal is very large, large enough to drive the grid to the positive saturation region. The plate-current excursion is thus made very large compared with the static plate current (which flows only during a small part of each signal cycle) and the efficiency is often

as high as 80 or 85 per cent. The wave forms in Fig. 191 show the great distortion produced in the plate current by this mode of operation, but the voltage appearing across the tuned plate circuit is nearly free from distortion, because of its low response to the high-frequency distortion components. To keep this distortion at a minimum, the tuned circuit employed as a plate load must be very "sharp," *i.e.*, highly responsive to the signal frequency and to no other. Consequently, the class *C* radio-frequency amplifier is not suited to the amplification of modulated-carrier signals, since it will discriminate between the carrier- and the side-band frequencies.

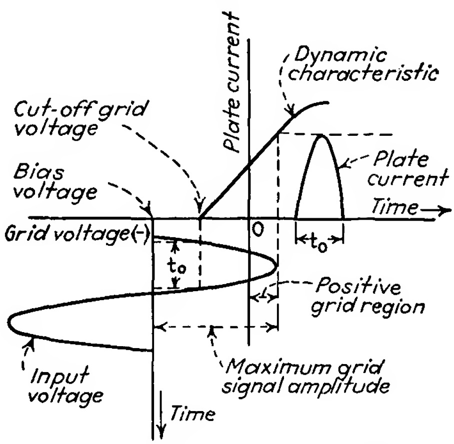


FIG. 191.—Class *C* dynamic operation.

67. Modulator Circuits.—The amplitude modulator as we have seen is a circuit to which two frequencies are fed: the high-frequency carrier and the audio-frequency modulating signal. The output of the modulator is an alternating current whose frequency is that of the carrier but whose amplitude depends on the instantaneous value of the modulating signal. Two general methods are available for combining the two frequencies. In the first, grid modulation, the radio and audio frequencies are both applied to the grid circuit of a tuned amplifier. The radio frequency generates the desired carrier frequency in the output of the tube, while the audio-frequency signal acts as a rapidly varying "bias" in the grid circuit, varying the power output of the tube in accordance with the modulating signal.

When large audio-frequency signals are applied to the grid circuit (thus producing a high percentage of modulation), some undesired distortion occurs, and the received (demodulated) signal will then have a wave form differing from that of the original modulating signal. Hence, faithful modulation at high percentages is difficult with grid modulation.

Much less undesired distortion occurs if the audio-frequency modulating signal is applied to the plate circuit, rather than to the grid. The grid is excited by a high frequency, thus establishing the carrier. If the tube is operated as a nonlinear device (as all modulators must be), the amplitude of the output of the circuit depends directly on the value of the applied plate voltage.

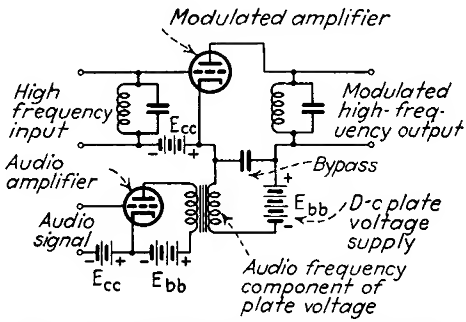


FIG. 192.—Plate-circuit amplitude modulator.

If the applied plate voltage is supplied in the form of a high-voltage audio-frequency signal, then the radio-frequency output of the tube is modulated in accordance with the audio frequency. While a powerful source of audio-frequency signal is required in this case, compared with grid modulation, the distortion is much reduced. Accordingly plate modulation (usually called Heising modulation after its inventor) is very widely used in broadcast transmitters, etc. A typical connection diagram of a plate-circuit modulator is shown in Fig. 192.

The modulated oscillator (used only in simple equipment, such as portable transmitters) is a self-excited oscillator whose plate voltage is varied by an audio-frequency amplifier. It is, in other words, a modulator which generates its own high-frequency carrier excitation. The chief disadvantage is the tendency of the self-excited oscillator to change its frequency as the plate voltage changes, hence the modulating signal pro-

duces changes in both the amplitude and the frequency of the carrier. The frequency modulation produces distortion unless the receiver tuned circuits have a very wide frequency response, and in any event it occupies more territory in the radio-frequency spectrum than simple amplitude modulation.

In plate-circuit modulation, when the percentage modulation is 100 per cent, the output current of the modulator is zero at the negative signal peaks and twice the normal carrier level at the positive signal peaks. This represents a variation in power from zero to four times the normal unmodulated power. All of the signal power is supplied to the modulator from the audio-frequency source, which must therefore be able to supply an *instantaneous* peak power of four times that of the normal unmodulated power consumed by the modulator.

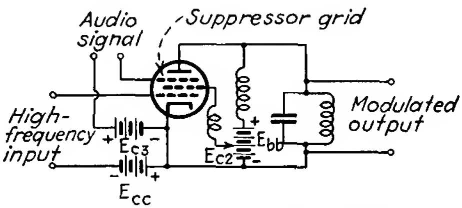


FIG. 193.—Suppressor-grid modulator.

A form of modulation used widely in transmitters designed for transmission both of code and of voice, is suppressor-grid modulation. In this circuit (Fig. 193) a pentode tube is used as an oscillator or tuned amplifier, the control grid being excited at the radio frequency of the desired carrier. The audio-frequency modulating signal is applied between the cathode and the suppressor grid of the tube. The action of the suppressor in changing the power output of the tube at an audio-frequency rate is similar to that in ordinary grid modulation, except that the radio and audio frequency are isolated from each other in the suppressor-grid scheme. The degree of modulation obtainable is 100 per cent, but the average power output, when so modulated, is only about 25 per cent as great as could be obtained from the same tube as a simple single-frequency class C amplifier, for telegraphy.

68. Demodulator Circuits (Detectors).—After the amplitude-modulated radio-frequency carrier is tuned and amplified at the receiver, it must be demodulated (Sec. 64, page 289) before the

original intelligence is reproduced. The demodulator tube is in general one which will respond more readily to positive peaks than to negative ones, and it is therefore a nonlinear device. Three important types of demodulators are now in use: the diode rectifier detector; the grid-leak grid-circuit detector; and the biased plate-circuit detector.

The diode is a very simple form, consisting simply of a vacuum-type diode and a load resistor connected to the source of the modulated carrier (Fig. 194(a)). The output current from the combination depends on the load line page 247 which is plotted

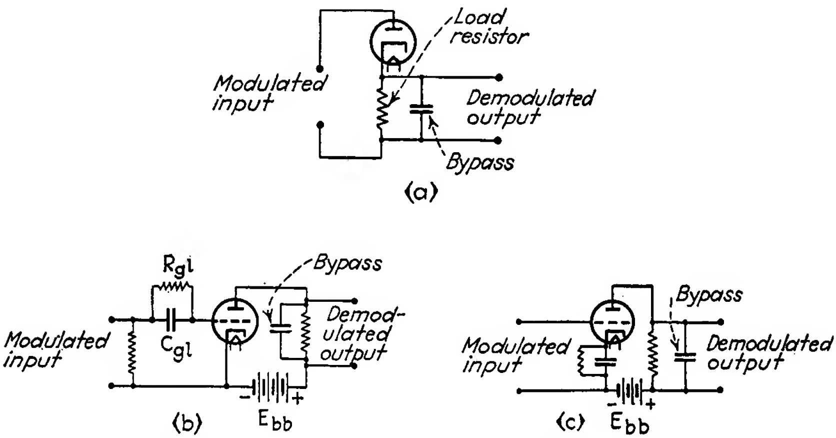


FIG. 194.—Demodulator (detector) circuits: (a) diode, (b) grid-leak, and (c) plate-circuit.

on the static characteristics in the usual manner. When a modulated-carrier voltage is applied, the output of the diode detector contains high-frequency ("carrier-remnant") components, and low-frequency components, the latter being the modulating signal and, in addition, distortion components. Also present is a direct current which does not change with the modulation, but which is proportional to the strength of the unmodulated carrier. This direct-current component is of no value as a signal, but it is of value for measurement and control purposes based on the strength of the received signal, especially for the automatic control of volume.

The *grid-leak detector* is a highly sensitive form, since it makes use of the amplification properties of the triode tube. The typical circuit diagram is shown in Fig. 194(b). The grid of the

tube is not biased, hence the positive peaks of the modulated-carrier voltage drive the grid positive, causing grid current to flow. This grid-current flow charges the grid capacitor C_{gl} to a value proportional to the peak value of the modulated radio frequency at each instant. The resulting voltage appearing across the capacitor biases the tube and influences the plate current accordingly. The grid-leak resistor R_{gl} is chosen so as to allow the charge across the capacitor to leak off rapidly enough so that the capacitor charge can change rapidly in accordance with the modulation variations in the incoming carrier. As a consequence the voltage across the capacitor is always changing in accordance with the modulation, and the voltage appearing across it is, therefore, a reproduction of the original audio-frequency modulating signal. The grid and cathode of the tube, in other words, act as a diode rectifier, the load being a shunt combination of resistance and capacitance. The signal across this combination, being applied to the grid circuit, is amplified by the tube and delivered to the plate load as an audio-frequency signal. Any high-frequency components are usually by-passed either in the grid circuit capacitor or in the plate-circuit by-pass capacitor.

The *biased plate-circuit detector* is a simple amplifier tube with its grid- and plate-bias voltage so chosen that the tube operates in a nonlinear portion of its static characteristics. The dynamic characteristic then emphasizes the upward peaks of the radio-frequency carrier relative to the downward peaks. Consequently the voltage across the plate-load impedance contains audio-frequency as well as radio-frequency components, the latter being by-passed to cathode. By choosing the load impedance properly, the degree of desired signal, relative to the distortion, may be made very large. The plate-circuit detector is, in fact, considered preferable to the grid-circuit type because of its lower distortion.

The action of all three types of detectors is very complicated if the complete performance under a wide variety of signal frequencies and modulation percentages is considered.

69. Multivibrator (Frequency-division) Circuits.—A multivibrator is a type of oscillator which does not use a tuned circuit (Fig. 195). Actually it closely resembles a two-stage resistance-capacitance coupled amplifier whose output is fed back to the

input. The feedback connection provides the requisite energy for sustained oscillations, since the second tube introduces a 180° phase shift. The frequency of oscillation depends primarily on the capacitance of the coupling condensers and the plate-load resistances used. If allowed to oscillate as a self-excited device the frequency is approximately

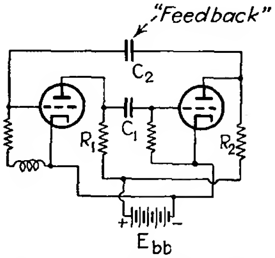


FIG. 195.—Multivibrator (frequency-division) circuit.

$$f = \frac{1}{R_1C_1 + R_2C_2}. \tag{94}$$

If a high-frequency generator is connected to the input terminals (Fig. 195), the output will contain many frequencies, several of which are submultiples of the exciting frequency. The circuit thus operates as a frequency divider; it is the only satisfactory method of accomplishing this function electronically. The ability of the circuit to generate submultiples is explained in terms of the "motorboat" action of the resistance-capacitance coupled amplifier, in which the periodic charging of the coupling condenser will block the amplifier for a period equal to several cycles of the exciting signal. Frequency division of 1 in 25 is readily accomplished.

70. Converter Detectors.—An important function employed in superheterodyne radio receivers is *frequency conversion* by a detector, commonly called the *first detector* or *mixer* tube. This

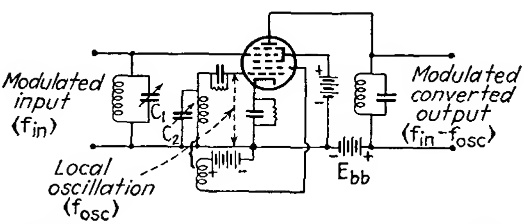


FIG. 196.—Pentagrid converter stage. One tube containing five grids performs the functions of oscillation, modulation, and amplification, in superheterodyne service. Capacitors C_1 and C_2 are varied simultaneously.

tube may be a simple plate-circuit detector, although more commonly a specialized pentagrid converter (Fig. 196) is used. To this tube are fed the modulated carrier voltage and another locally generated alternating current of a frequency separated from the carrier frequency by a fixed frequency interval, usually

between 400 and 500 kc. The nonlinear action of the detector tube combines or "mixes" these signals, and at the output of the detector there appears a new modulated carrier whose frequency is equal to the interval between the original carrier and the locally generated radio frequency. As the circuit is tuned, the frequency of the locally generated alternating current is changed also so that the outgoing frequency ("intermediate frequency") always is the same, namely, that of the interval between the original carrier and the locally generated alternating current. This intermediate frequency is then amplified by tuned amplifiers which are fixed in tuning, since the output of the mixer tube is fixed in frequency. Since no variable tuning of these amplifiers is necessary, they may be adjusted once and for all for maximum gain consistent with the necessary side-band response. The signal thus receives the same amplification regardless of its original carrier frequency, *i.e.*, the gain is the same at all points of the tuning dial. Furthermore, the intermediate frequency is chosen to have a value at which amplification by tuned amplifiers is both stable and of high gain. The advantages of this method of reception have made the simpler tuned radio-frequency receivers (in which each radio-frequency stage is separately tuned) almost obsolete. The intermediate frequency, being a modulated carrier, must then be demodulated in a second detector in order to realize the audio-frequency signal components, which are then amplified and applied to the loud-speaker.

71. Automatic Circuit Control.—The amplification provided by a circuit may be regulated by the use of voltage dividers, or by varying the amplification factor of the amplifier tube directly by adjusting the negative grid-bias voltage applied to it. The latter method is commonly used in radio receivers. It permits automatic volume control (a.v.c.) by the scheme shown in Fig. 197. The output of the a-v-c diode tube contains a direct current whose value is proportional to the strength of the incoming carrier. If this direct current is passed through a resistance and the resulting voltage drop applied as a bias voltage to the grid of the preceding amplifier tube, then the amplification provided by this tube will be adjusted in direct proportion to the strength of the incoming carrier, and in such a way that a strong signal will receive little amplification, a weak one much amplification. To accommodate a wide range of automatic-

volume-control action, it is necessary that the controlled amplifier tubes be capable of operating without distortion under a wide variety of grid-bias voltages. The supercontrol tubes (variable μ types) are especially designed with this fact in mind and are universally used in a-v-c service.

Tubes have also been used for adjusting the audio-frequency response (tone) automatically in terms of the strength of the incoming signal, and for adjusting the frequency of a superheterodyne oscillator so that its frequency interval from the incoming carrier is always at the correct value. The latter

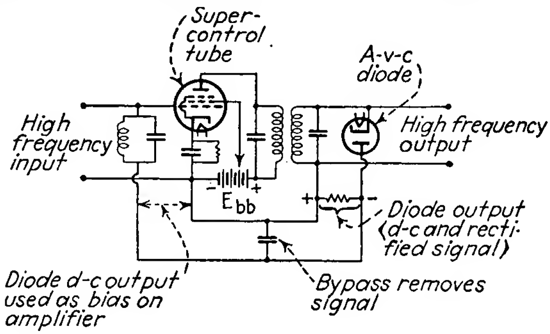


FIG. 197.—Automatic volume control.

circuit (automatic frequency control) makes misadjustment of the tuned circuits of a receiver much less likely than when no automatic control is used, and is consequently used in connection with "push-button" tuning and remote-control tuning devices in which precise tuning is difficult to accomplish.

Problems

1. Express the following power ratios in decibels: 65 to 1, 250 to 1, 1 to 1000, 1,750,000 to 1; and the following voltage ratios in decibels voltage gain: 16 to 1, 1 to 256, 10,000 to 1, 32 to 145.

2. A push-pull class *B* amplifier (Fig. 177) is fed a sine-wave signal of 16 volts per tube, maximum amplitude, and is biased at cutoff (-14.0 volts). The load resistance is 20,000 ohms per tube, and the plate voltage 250 volts. Using the static characteristics of Fig. 152, compute the power output by the load-line method, per tube, and the efficiency of the plate circuit (power developed in load resistance divided by average plate power). Compare this efficiency with that of the class *A* amplifier computed in Prob. 6, Chap. XI, page 252.

3. Compute the gain (e_{g2}/e_{g1}) of the resistance-capacitance-coupled amplifier of Fig. 180, at 100, 1,000, and 10,000 cycles per second. $g_m = 1500$ μmhos; $r_p = 10{,}000$; $R_p = 10{,}000$; $R_g = 200{,}000$; $C_c = 0.01\ \mu$f.; $C_s = 100\ \mu\mu$f.

4. A carrier voltage of 1,000,000 c.p.s. frequency and 100 volts maximum amplitude is modulated by an audio-frequency signal of 1000 c.p.s. frequency and amplitude 50 volts. Compute the percentage modulation and the frequencies of the side bands. Sketch roughly the amplitude-versus-time plot of the resulting modulated carrier.

5. The modulated carrier of Prob. 4 decreases in strength by 20-decibel voltage loss in being transmitted to a diode detector (static characteristics, Fig. 57) whose load-resistance is 10,000 ohms. By graphical means find the maximum amplitude current of the demodulated signal frequency in the load resistance, and the value of the direct-current component.

Bibliography

ALBERT, A. L.: "Electrical Communication," John Wiley & Sons, Inc., New York, 1934.

HENNEY, KEITH: "Principles of Radio," John Wiley & Sons, Inc., New York, 1936.

HENNEY, KEITH and others: "Radio Engineering Handbook," McGraw-Hill Book Company, Inc., New York, 1935, Secs. 9, 10, 11, 12.

PENDER and McILWAIN: "Electrical Engineering Handbook," vol. V, Communications, John Wiley & Sons, Inc., New York.

HUND, A.: "Phenomena in High Frequency Systems," McGraw-Hill Book Company, Inc., New York, 1936.

HUND, A.: "High Frequency Measurements," McGraw-Hill Book Company, Inc., New York, 1933.

CHAPTER XIV

INDUSTRIAL-CONTROL AND MEASUREMENT CIRCUITS

Introduction.—The utility of the electron tube in industrial-control and measurement problems is based on four basic abilities: (1) the ability to amplify weak control impulses; (2) the ability to react practically instantaneously; (3) the ability to react to the stimulus of light; and (4) the ability to measure minute currents and voltages without appreciably disturbing the power source. These divisions in themselves do not reveal the great variety of specific applications and combinations of functions which can be made. Often a single electronic relay, for example, will combine sensitivity with light, high amplification and great rapidity of response. In general, the electronic circuits used for control and measurement are not so easy to classify as the communications circuits, since the industrial devices are designed usually with some special problem in mind, and the forms they take are nearly as numerous as the problems themselves. It is possible, therefore, to select only a few circuits which have broad usefulness and to indicate the several ways in which they are used.

72. Electronic-relay Circuits.—The broadest class of electronic-control circuits are the electronic relays, in which an amplifier tube serves to transfer a control impulse in amplified form to an electromechanical relay. The control impulse is applied in general between grid and cathode while the plate current flows through the relay coil. When the net voltage between grid and cathode changes by a sufficient amount, the plate current passes through the "pull-up" or "drop-out" values of the relay which closes or opens accordingly. The tube serves a purpose, of course, only when the control voltage is not sufficiently great to operate the relay directly, or when it is desired to operate the relay without drawing current from the control source. An example of the latter case is the simple contact relay used for humidity and temperature control. In

temperature-control work a very small thermostat is sometimes used. The delicate contacts of this thermostat must control a relay which in turn controls the heating power to be regulated. If the contacts of the thermostat are connected directly in series with a relay coil and battery, the sparking at the contact points will in time oxidize the contact surfaces, increasing the resistance in the coil circuit to the point where the relay will not close. If the arcing at the contacts is severe, the contacts may stick together by the welding action of the current flow. If, however, the contacts are used simply to apply a voltage to the grid of a vacuum-tube amplifier, in whose plate circuit the relay is connected (Fig. 198), then the current flow across the contacts may be made so small that no arcing whatever occurs and the contacts will remain clean and in working order indefinitely. The combination of the tube and the relay is, in other words, a voltage-operated device. It is used whenever the control contact points are necessarily small and cannot carry the required relay current without eventual failure. The absence of arcing is also of use when the control contacts must work in an explosive atmosphere.

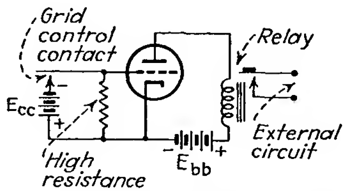


FIG. 198.—Grid-contact relay circuit.

Time-delay Relays.—When a capacitor is charged from a direct-current source through a series resistance, the voltage across the capacitor increases at a rate which depends on the size of the capacitor and the series resistance. If the capacitor and resistance values are large, the time required to build up an appreciable voltage may be many minutes. This fact gives rise to a very simple method of introducing a time delay in electric circuits. If the voltage across the capacitor is used to operate a relay, the relay will close only after a definite time has elapsed after the charging begins. Unfortunately this cannot be accomplished directly because the low resistance of the relay coil across the capacitor discharges it more quickly than it can charge through the high-value series resistance. If an

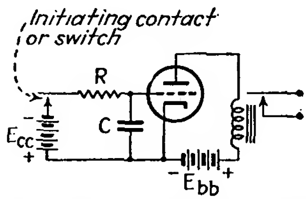


FIG. 199.—Time-delay relay circuit.

amplifier tube is connected across the condenser (Fig. 199), properly biased so that its grid does not draw current, then no discharge current can flow, and the capacitor voltage then builds up at the required rate. The relay, placed in the plate circuit of the tube, opens when the capacitor voltage has reached a value corresponding to the plate current required to open the relay.

A very wide variety of different delay circuits have been built up around this basic idea. If, for example, a set of contacts on the plate-circuit relay is caused to discharge the capacitor when the relay opens, the circuit can be made to repeat its action at regular intervals, *i.e.*, to "count time." The time interval may be adjusted by properly proportioning the capacitor and resistor values. If the charging voltage is E volts, the series resistance is R ohms, and the capacitance is C farads, then at the end of t sec. after the charge begins, the voltage E_c across the capacitor is

$$E_c = E(1 - \epsilon^{-t/RC}) \text{ volts.} \tag{95}$$

A practical rule, derived from this equation is: At the end of $t = RC$ sec. the capacitor voltage E_c is 63.2 per cent of the charging voltage E. The *discharge* of a condenser from its fully charged condition may also be used. In this case the equation is

$$E_c = E(\epsilon^{-t/RC}) \text{ volts} \tag{96}$$

where E is the original voltage across the condenser before the discharge begins.

The condenser-charge circuit may also be used to measure current impulses. If a current of I amp. flows for t sec. into a condenser of C farads, the voltage change ΔE_c across the condenser is

$$\Delta E_c = I \times \frac{t}{C} \text{ volts.} \tag{97}$$

This change in voltage may be measured by noting the corresponding change in plate current, as indicated by a meter in the plate circuit. The same method may be used to store several current impulses (current integration) until the net voltage change in the grid circuit, due to the combined action

of several stored pulses, is sufficient to produce the required closing current in the control relay.

Capacity-operated Relays.—A somewhat more complicated type of electronic relay, of great practical usefulness, is the capacity-controlled relay, in which the grid voltage of the relay tube is controlled by a small change in electrical capacitance between two electrodes. Usually a rather elaborate control circuit is necessary to provide a sufficiently great change in the control voltage from the necessarily small capacitance changes. One typical circuit which illustrates many circuit principles is the capacity-operated relay of Shepard, shown in Fig. 200. The tube A is an oscillator, arranged in a Hartley circuit, with an antenna (which serves as one of the electrodes) connected

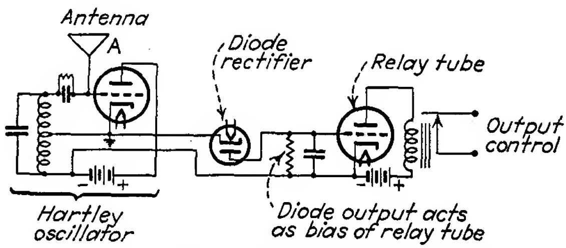


FIG. 200.—Capacitance-operated relay circuit.

to the grid. Any increase in the capacitance between this antenna and the ground (cathode) will increase the vigor of oscillation, and hence increase the amplitude of the high-frequency alternating current generated across the tuned circuit. Part of this alternating voltage is rectified in a diode rectifier and the diode current used to charge a resistor-capacitor combination. The voltage appearing across the capacitor is then proportional to the oscillating voltage, which in turn is controlled by the capacitance between the antenna and ground. The diode load voltage is applied as the control voltage of the relay tube, in whose plate circuit the final relay operates. The sensitivity of this arrangement can be made very great. The motion of one's hand 4 ft. from the antenna, a capacitance change of a very small fraction of a micromicrofarad, is sufficient to operate the final control relay. The suitability of the device as a burglar alarm is readily apparent. Suitably modified it may be used as a thickness gauge, for measuring the thickness

and the moisture content of paper or textiles, and for controlling them in manufacturing processes.

Phototube-controlled Relays.—The type of electronic relay capable of the widest diversity of applications is undoubtedly the phototube-controlled relay. In this type of relay the control voltage in the relay-tube grid circuit is obtained from a phototube current passed through a grid-coupling resistance. Two types of connection, positive and negative, can be used. In the positive circuits an increase of light on the phototube causes an increase

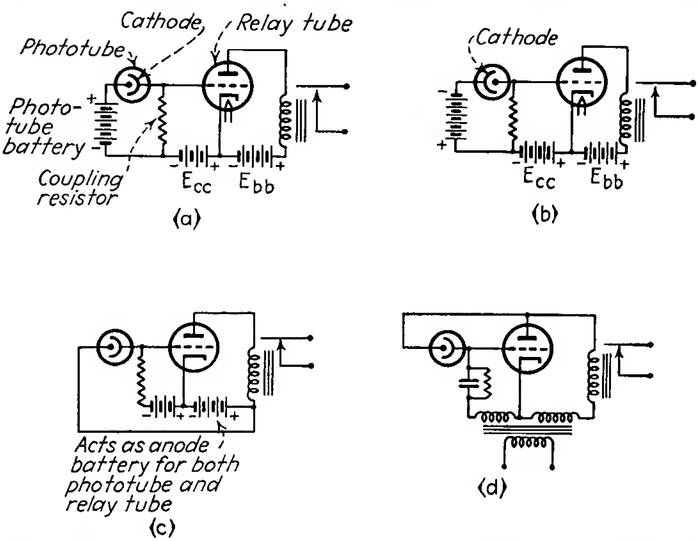


FIG. 201.—Phototube relay circuits: (*a*) positive, (*b*) negative, (*c*) common-anode battery, positive, (*d*) alternating-current operated.

in the plate current of the relay tube; the cathode of the phototube is connected to the grid of the relay tube, as shown in Fig. 201(*a*). The negative circuits, producing a decrease in plate current with increase in light, use the phototube with its polarity reversed, *i.e.*, with its anode connected to the grid of the relay tube. These circuits require a separate phototube battery in the grid circuit to serve as a source of the phototube current. In the positive type of circuit it is possible to avoid the use of this extra battery by connecting the phototube between grid and the positive terminal of the relay-tube anode battery. This battery thus does double service, providing both relay-tube plate current and phototube current (see Fig. 201 (*c*)).

To calculate the action of any of these circuits it is necessary to use a load line on the current-voltage characteristic of the phototube, as shown in Figs. 97 and 98 (page 165). This load line shows the voltage appearing across the tube at any light level, the difference between this value and the net applied voltage being the voltage appearing across the grid-coupling resistor. This latter voltage acts as the "signal" in controlling the relay-tube plate current.

For many purposes the use of direct current for relay- and phototube currents is inconvenient since batteries or a source of rectified alternating current must be supplied. An alternating-current-operated phototube relay, shown in Fig. 201(*d*), overcomes this objection by applying alternating current to both plate and grid circuits. A resistor-condenser combination is used in the grid circuit to shift the phase of the grid-circuit alternating current with respect to that of the plate circuit. The current flowing in the plate circuit is controlled by the action of the phototube (connected between grid and plate) in altering both the amplitude and the phase of the grid voltage with respect to the plate voltage, as the light falling on the phototube changes in intensity.

Uses of Phototube Relays. Illuminating Control.—The adaptability of the phototube relay may be best illustrated by several examples of its use in industry and general control practice. Among the most important are the illumination controls. The phototube relay is set up so that the phototube views a "control surface" such as a work bench, school desk, warehouse wall, or some other lighted surface, the illumination of which is to be controlled. The relay in the output of the phototube amplifier is connected, either directly or through subsidiary relays, to a source of artificial illumination. As the illumination of the control surface decreases (owing to the approach of nightfall, the presence of a storm, or for any other reason) the phototube current decreases and the relay finally pulls up, connecting the source of artificial light to counteract the loss of natural illumination. The level of illumination at which this action occurs depends, of course, on the application. For warehouse lighting the level is very low, for schoolrooms the level should be high enough to protect the eyes of the students. The circuit can be adjusted, of course, by varying the value of the grid-coupling

resistor, the bias applied to the relay tube, or by adjusting the mechanical tension of the plate-circuit relay. Often two or more relays are used, set for different levels, so that additional numbers of lights may be turned on as they are needed. The same general principle is used for turning advertising signs and street lights on at nightfall and off at daybreak; for measuring the density of smoke or the turbidity of liquids by passing light through them to the phototube, and for hundreds of other similar purposes.

Counting and Sorting.—A different class of phototube relays are the counters, graders, and sorters. In these applications, light is not present inherently as it is in illumination problems but must be supplied from a lamp or other source. In the

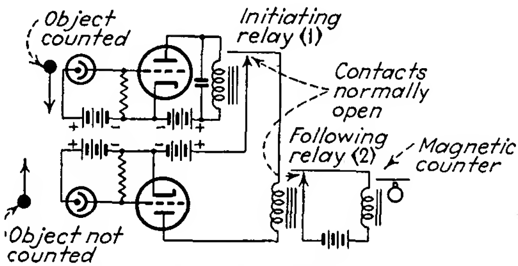


FIG. 202.—One-way phototube counting relay.

simplest counting arrangement the source of light entering the phototube is interrupted by the motion of the object to be counted, and the corresponding change in phototube current operates a magnetic counter in the plate circuit of the relay tube. This device may be used for counting traffic on a highway, objects on a conveyor belt or chute, and in general wherever the counted objects are in motion past a fixed point. Variations of the simple counter can readily be made; in Fig. 202 is shown a nonreversible counter which will count objects moving past it in one direction but not in the other. The interconnections between the relays in the plate circuits of the two relay tubes are arranged so that the counting circuit is completed in the 1-2 sequence, but not in the 2-1 sequence. The speed of counting is limited only by the ability of the magnetic counter to record the impulses; one circuit (Fig. 203) uses gas-triode tubes to store impulses in groups of 2, 4, 8, etc., so that the rate of the magnetic counter may be slower than the actual rate of counting. Up to 5000 impulses

per second may be registered in such a system. The advantage of the phototube method of counting is that the counted objects are not interfered with in any way except by the passage of the intangible beam of light.

The graders and sorters may be simple or complex, depending on the number of grades into which the objects are to be sorted.

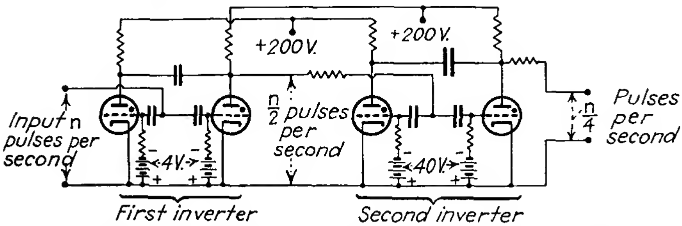


FIG. 203.—Impulse divider circuit for high-speed counting.

In the simple amplitude sorter, objects larger than a certain size are sorted from objects smaller than this size. The beam of light passing the objects falls into one of two phototubes, depending on the height of the object. The control relays connected to each phototube operate a magnetic sorter which deposits the

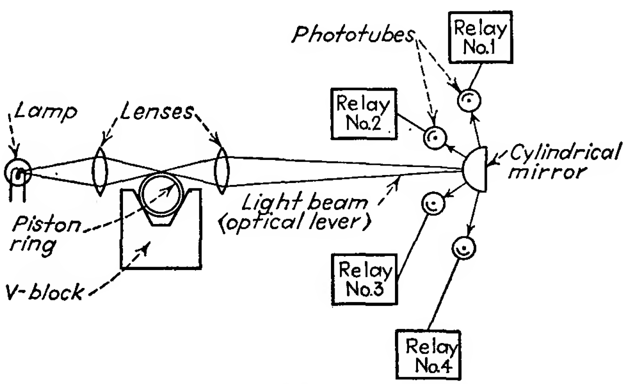


FIG. 204.—Optical size-measuring device employing phototube relays.

objects in one of two bins, depending on their size. Magnetic counters may be used in conjunction with the sorters to show the number of objects in each size classification.

The same principle may be applied to sort objects into many sizes. A piston-ring sorter developed by Powers is shown in Fig. 204. The piston ring, resting on a V-block, allows light to

fall into one or more phototubes, depending on the diameter of the ring. The use of a long optical lever gives a high degree of precision, so that the differences in diameter between the excitation of the different phototubes is only 0.0001 in. The use of the hemicylindrical mirror directs the light into phototube 1 if the piston rings are of large diameter, into phototubes 1 and 2 if the diameter is the next size smaller, and so on. The relays connected to the phototubes operate a four-place magnetic sorter which deposits the rings in the proper bins.

Sorting of objects may also be made on the basic of color, rather than size. In this case two or more phototubes are fitted with color filters, so that they react only to objects of the color of the filter. By arranging the several phototubes so that they view the passing objects, the relay corresponding to the color of the object will go into action and operate a magnetic sorter which in turn places that object in the proper bin. Phototubes of different spectral response may be used to accentuate the difference in color sensitivity, and double-filter arrangements may be used in connection with two phototubes and a single sorter to separate any shade of one color from any shade of another.

Phototube Door-openers, Speed-traps, Etc.—The simple phototube relay may be used for opening doors automatically. The person or vehicle approaching the door interrupts a light-beam passing from the light source to the phototube. The relay in turn operates a magnetic valve on a pneumatic plunger connected to the door. Such installations are highly useful in restaurants where the tray-laden waiters cannot use their hands to open the door, and in public places, such as railroad stations, where pedestrian traffic is heavy. It has also been applied to the automatic opening of garage doors in commercial and private fields.

A more complicated use of the phototube relay is the speed trap, which operates an alarm when a vehicle passes along a road at an excessive speed. Two phototubes and corresponding light sources are set up beside the road, separated by a definitely known distance. The car, passing the beams, interrupts them consecutively and so operates their associated relays. These relays in turn initiate and terminate the operation of a timing circuit (which may be of the condenser-charge variety discussed

above). If the time between the consecutive relay operations is very short (corresponding to a high rate of speed of the vehicle) an auxiliary relay operates to close the alarm circuit. On the other hand, if the interval between relay operations is long (slow speed) no alarm is given. By measuring the charge accumulated across the condenser in the timing circuit it is possible to measure the speeds of the passing cars automatically. The principle is not restricted to traffic problems, of course, and may be applied to a wide variety of velocity-measurement problems in industry.

Meter-and-mirror Phototube Relays.—Whenever a vacuum tube is used as a relay device, the control impulse must be in the form of a voltage change, since the tube is a voltage-operated device. But in many control problems the control impulse comes from a low-impedance source, and as a consequence the voltage developed from it is very small, often too small to initiate the control action. One very important case of this kind is the measurement of temperature with a thermocouple, a highly desirable method since it provides a continuous electric current proportional to the measured temperature. The impedance of the thermocouple is low, usually not more than a few ohms, and the current flow usually not more than a few milliamperes. The voltage output from a thermocouple is therefore a few millivolts, and the change in voltage corresponding to a small change in temperature may be but a few microvolts, far too small to control a relay tube. If the thermocouple current is passed through a current meter of appropriate sensitivity, however, the pointer will move and its indication will measure the temperature. The power available is far too small to operate a relay for control purposes, so the problem arises of coupling the pointer to a relay indirectly. The solution is the meter-and-mirror phototube relay. On the pointer of the meter is cemented a small mirror, which can reflect light from a lamp into a phototube. When the pointer reaches a predetermined reading, the light enters the phototube; at higher or lower readings the tube does not receive the light. The relay connected to the phototube thus operates when the meter pointer reaches the control position. Essentially the meter-and-mirror combination converts the low-voltage control impulse into a change in phototube current which can develop a large voltage change in the grid circuit of the relay tube.

Many variations of this basic idea have been evolved. One of the simplest is the hole-in-the-meter method. A small hole is drilled in the face of the meter, a light source placed behind it and the phototube in front. When the meter pointer rises to a sufficient value to cover the hole, the phototube current drops and the control relay comes into action. Another form is designed to prevent overshooting. If the meter pointer accidentally moves higher than the control position, the control is no longer effective and the heater circuit will then "run away." But if two holes are used in the meter face, with two phototubes, their relays can be connected to the heat source so as to keep the

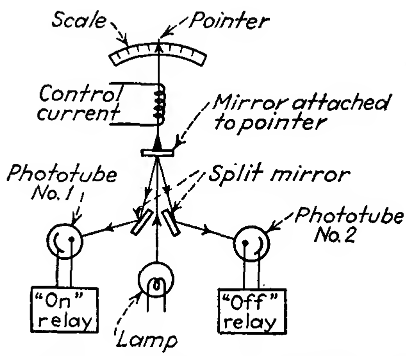


FIG. 205.—Meter-and-mirror on-off relay for coupling a low-impedance control circuit.

pointer within the control range between the two holes. A similar arrangement for the same purpose is the double-mirror arrangement. Light is reflected from the mirror on the pointer to two other mirrors arranged in roof-like fashion, as shown in Fig. 205. If the meter pointer is above the control position, the light enters phototube 1, otherwise it enters phototube 2. The relays associated with the two phototubes turn the heat source off and on, respectively.

Gas-filled Versus Vacuum Tubes in Relay Service.—The relay tubes referred to in the foregoing paragraphs may, in general, be of any type: triode, tetrode, or pentode, gas-filled or highly evacuated. The vacuum tubes have the advantages of high speed of action, negligibly small grid current when properly biased, and continuous grid control over the entire operating range. They are especially suited to measurement purposes,

i.e., those in which the relay must operate only when a definite quantitative level of the control factor has been reached. The vacuum tubes are, on the other hand, considerably limited in the amount of plate current they can supply to the relay, and hence relays used with them must be of sensitive, and therefore delicate, design. When a large amount of power is to be controlled, the sensitive relay must be used to control the coil current of a larger power-handling relay. This is an expensive procedure, and one which introduces an additional possible source of failure.

The gas-filled tube, on the other hand, can supply large amounts of power in the plate circuit, not only because the current-carrying capacity of these tubes is large, but also because the voltage drop in the tube itself is small, leaving a large amount of voltage available for closing the relay during the conduction period. The grid current of the gas-filled thyratron, on the other hand, is much larger than that of the vacuum tube, regardless of the bias conditions, and the control of the grid over the plate current is far from continuous. This latter disadvantage makes it necessary either to operate the plate circuit on alternating current or to provide some other means of reducing the plate voltage below the extinction value once during each control cycle. This discontinuous control also limits the utility of gas-filled tubes in measurement service. But thyratrons are used whenever possible because they can control a very rugged relay directly and may in fact carry the full-load current of the controlled device, thus eliminating the need for any electromechanical relay whatever.

73. Electronic Welding-control Circuits.—Resistance welding is performed by passing a very heavy current through the pieces of metal to be joined. The heat generated in the resistance of the contact is sufficient to bring the work to the melting point, whereupon pressure on the joint forms the weld. If the current is excessive or if it lasts too long, the heat generated will exceed the melting-point value, resulting in oxidation and crystal-structure changes which weaken the weld. On the other hand, if the current is too weak or lasts for too short a time, the heat generated will be insufficient to fuse the metal surfaces. For these reasons very accurate control of both the amplitude and duration of the current flow are required to produce successful welds. The problem is accentuated by the necessity of high-

speed operation with rapid and accurate welding cycles occurring as often as 100 or more per minute. When a weld is to be made along a seam between two sheets, for example, continuous welding is difficult, so it is customary to make separate spot welds along the seam, each overlapping the next. To produce uniform strength in such welding, a very high degree of accuracy in current control is required.

Gas-filled electron tubes are ideally suited to this type of service, since they can carry large currents under accurate control. The ignitron type of tube is especially well suited because of its high current capacity (peak currents of up to 5000 amp.). Currents of even this high value are not in themselves sufficient

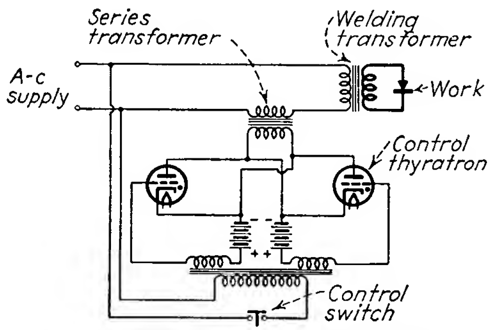


FIG. 206.—Indirect welding-control circuit.

to produce large welds, so it is universal practice to employ the current-control tube in the primary of a current step-up welding transformer in whose secondary currents of many hundreds of thousands of amperes can flow.

Two types of welding control are used: the direct and the indirect circuits. A typical example of the indirect circuit is shown in Fig. 206. The primary current of the welding transformer passes through the primary of a control transformer in whose secondary are connected the two thyratron or ignitron tubes which control the flow. The tubes are connected "back-to-back," so current flows on both negative and positive half cycles. When the control switch is closed, the thyratrons conduct, short-circuiting the secondary of the control transformer. The primary impedance of this transformer thereby is caused to drop to a very low value, passing a surge of current to the welding transformer and, through it, to the work.

The direct type of circuit contains the control tubes directly in series with the welding transformer. Since the full primary current of the welding transformer must pass through the tubes, their current capacity must be great; hence ignitrons are usually used. In Fig. 207, the ignitrons are shown connected as usual in back-to-back fashion. The control of the ignitor circuit in each ignitron is taken by the auxiliary thyratron tubes which supply the ignitor current at the proper time. The current in leading control tube is initiated by the control switch in its grid circuit. When this switch is closed and when the anode of the tube is positive, the first current pulse passes. On the next half cycle the trailing ignitron must fire, and it does so only

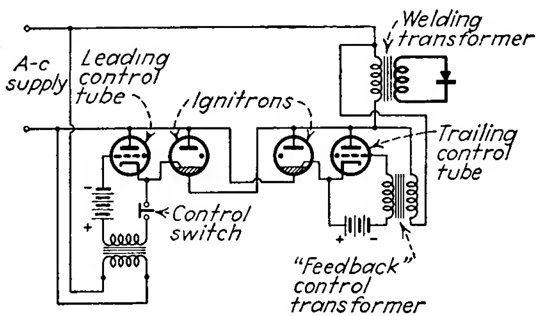


FIG. 207.—Direct welding-control circuit.

when its control tube becomes conducting. The control impulse for this tube comes from a feedback-control transformer connected directly across the welding transformer primary. The inverse surge of current passing to the welding transformer is thus fed back in part to the trailing control tube, which thereupon fires the ignitron. The control of the duration of each current surge is obtained by biasing the grids of the leading and trailing thyratrons. When the control switch is open, the surges cease at the end of the cycle. High-speed operation may be obtained by a mechanical contactor for closing the control switch as often as is desired.

74. Voltage- and Current-regulation Circuits.—Electron tubes are adapted to voltage- and current-regulation service because they react to small changes in current or voltage and are capable of producing large compensation in their plate circuits. A typical voltage regulator, suitable for low-current applications,

is shown in Fig. 208. It employs a vacuum-type triode, a voltage divider in the grid circuit, a bias battery, and a compensation resistance in the anode circuit. Assume a change of voltage across the input terminals of ΔE_i volts. The voltage change applied by the voltage divider to the grid circuit will be

$$\Delta E_g = \Delta E_i \frac{R_1}{R_1 + R_2}.$$

This change in grid voltage will produce a change in plate current of approximately

$$\Delta I_p = g_m \Delta E_g = \Delta E_i \frac{g_m R_1}{R_1 + R_2},$$

neglecting the effect of the change in plate voltage on the g_m value. This change in plate current passing through the compensation

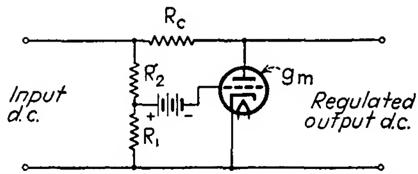


FIG. 208.—Voltage regulator for low-current applications.

resistor R_c will produce a change in voltage drop of

$$\Delta E_c = R_c \Delta I_p = \Delta E_i \frac{g_m R_c R_1}{R_1 + R_2}.$$

Now, if the compensation-voltage change ΔE_c equals the original change in voltage ΔE_i, then the voltage across the cathode and plate of the tube, the output voltage, will remain unchanged. Hence the condition for voltage regulation is, from the above equation, since $\Delta E_c = \Delta E_i$,

$$\frac{g_m R_c R_1}{R_1 + R_2} = 1. \tag{98}$$

By properly proportioning the resistance and mutual-conductance values, therefore, the output voltage can be made to remain constant over a wide range of changes in the input voltage.

When a large amount of voltage-regulated power is required, the simple scheme just described is not suitable. Several other

methods are available, however. In one of them (see Fig. 209), a grid-controlled rectifier is used to supply the field current to an alternating-current (or equally well to a direct-current) generator. The output of the generator is connected to the grid circuit of the controlled rectifier in the proper polarity, so that any decrease in the output voltage of the generator causes an increase in the current flow from the rectifier to the field winding. This increase in field current compensates for the original decrease in voltage. Another system makes use of a reactance coil in series with an alternating-voltage supply. A separate winding on the same core receives current from a grid-controlled rectifier whose grid circuit is controlled by the input voltage of the circuit. Any decrease in this voltage causes the rectifier

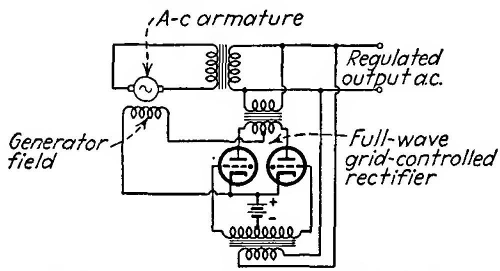


FIG. 209.—Alternating-voltage regulator employing grid-controlled rectifier.

to pass more current through the auxiliary winding which, by saturating the magnetic core of the reactance, decreases its reactance value. The resulting decrease in voltage drop compensates for the original decrease in voltage, so that output voltage remains constant.

The same method may be applied for current regulation by insuring a constant voltage value across the terminals of the current-carrying source. Current regulation may also be performed by any tube with a very flat current-voltage characteristic, such as a saturating diode tube or a pentode tube.

75. Electronic-measurement Circuits. Vacuum-tube Voltmeters.—The ordinary direct-current voltmeter, of the D'Arsonval moving-coil type, is in reality a current-operated device, since it depends upon a current flow to create the change in the magnetic field on which the movement of its pointer depends. This type of meter therefore always consumes power from the circuit being measured, the power being equal to the product of the voltage

across its terminals times the current flow. In sensitive meters the current flow is small and, in low-resistance circuits, can be supplied readily without causing any appreciable error in the measured voltage. In high-impedance circuits, on the other hand, the current flow causes an appreciable increase in voltage drop in the circuit and a consequent error in the measurement. To avoid this difficulty a voltmeter is required which consumes no input current. The vacuum-tube voltmeter serves this purpose since the grid circuit, when properly biased, consumes no grid current. The voltage appearing across grid and cathode is transformed by the tube into a corresponding current reading

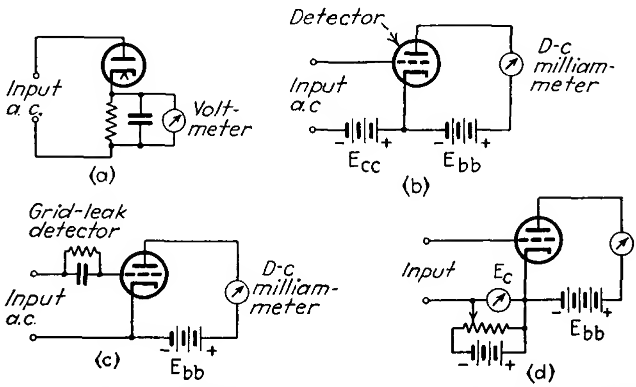


FIG. 210.—Vacuum-tube voltmeters: (*a*) diode, (*b*) plate-circuit, (*c*) grid-leak, and (*d*) substitution (slide-back).

in the plate circuit which may be measured directly by a milliammeter. This type (called a plate-circuit meter) is illustrated in Fig. 210*b*. Its disadvantages are limited range (since the applied voltage cannot exceed the bias value without causing grid current to flow) and the change in calibration which occurs with the aging of the tube. It is widely used, however, for measuring small voltages in high-impedance circuits.

Vacuum-tube Voltmeters for High-frequency Alternating-current Measurements.—Ordinary alternating-current voltmeters of the copper oxide rectifier or moving-vane types are severely limited in frequency range. When alternating current of higher than 5000 c.p.s. frequency is to be measured, usually some sort of vacuum-tube voltmeter is used. One type is the diode rectifier meter (Fig. 210*a*). The alternating voltage is applied to the

diode and is rectified by it. The output current contains a direct-current component and several alternating-current components. If the input alternating current has a sine-wave shape, the direct-current component bears a direct relationship to the amplitude of the input wave, and consequently a direct-current meter in the load circuit of the diode can be used to indicate this amplitude. For sensitive measurements it is sometimes desirable to amplify the direct-current component in a direct-current amplifier (see below) before measurement in a current meter.

The *grid-circuit triode voltmeter* serves the same purpose by its detection action. The grid-cathode circuit serves as a rectifier, building up a direct-current component across the grid leak and condenser (Fig. 210*c*), which in turn causes a measurable change in the plate current. This type of meter, combining both rectification and amplification in a single tube, is very sensitive. Both the diode and grid-circuit triode meters consume current from the circuit under measurement, but the current is usually small.

Measurements of the wave form of high-frequency alternating voltages can be made by means of amplifiers feeding a cathode-ray tube, as outlined in Chap. X.

Electronic Electrometers.—The same technique as used in the plate-circuit voltmeter may be used to measure extremely small currents, provided that a tube whose grid current is very small is used. Special electrometer tubes (notably the General Electric *FP*-54 and the Westinghouse *RH*-507) have been specially designed for this service. They are characterized by low-amplification constant (that of the *RH*-507 is less than unity), low values of mutual conductance, less than 100 micromhos, and by extremely small control-grid current when properly biased. The *FP*-54, for example, when biased to -4.0 volts draws 10^{-15} amp. grid current. The need for low grid current is illustrated by the circuit in Fig. 211. The input current to be measured flows to the grid resistor and the grid of the tube. If the grid current is small, the current will produce a voltage drop across the grid resistance which is accurately proportional to the input current itself; otherwise the grid current will subtract from, or add to, the input current. The tube transforms the voltage drop into a plate-current reading which is indicated by the microammeter or galvanometer in the plate circuit. A

simple circuit such as that of Fig. 211 is capable of measuring currents of 10^{-14} amp. (one-hundredth of a micromicroampere). It has been used, for example, to measure the phototube currents resulting from the light of a star, and in similar astronomical problems.

Direct-coupled Voltage and Current Amplifiers.—The direct-coupled amplifier (Sec. 63) is used in measurements work whenever a weak direct current or voltage or very slowly varying alternating current or voltage is to be amplified. A satisfactory direct-coupled amplifier for this purpose is very difficult to construct and operate since the absolute magnitude of the output current depends not only on the direct-current signal but also on the bias values applied to the tube, both of which are direct-current components and indistinguishable from one another.

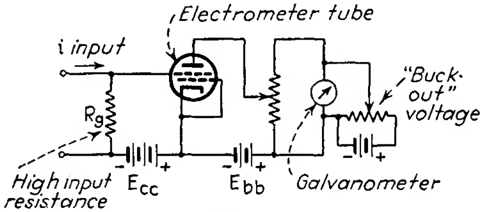


FIG. 211.—Electrometer circuit.

If the tube characteristics change or if the bias voltages vary, the accuracy of the measurement is greatly impaired. One source of difficulty is the variation in the "contact potential" between the grid and cathode of the tube. This contact potential results from the difference in the electrochemical properties in the metals used for grid and cathode structures; it depends to a considerable extent on the treatment of the metals during manufacture and upon subsequent changes as the tube ages. The potential acts as part of the bias applied to the grid and hence is, like the bias, indistinguishable from the voltage to be amplified.

One avenue of approach reduces these effects by employing two tubes in a circuit arranged to balance out changes in plate voltage and in tube characteristics, provided that the changes occur equally in both tubes. This type of amplifier, originally developed by Wynn-Williams, is in the form of a Wheatstone bridge, as shown in Fig. 212. The ratio-arm resistors are equal in value and the two tubes, which form the remaining two arms

of the bridge, are biased to offer the same resistance. When the signal is applied in the grid circuit of one of the tubes, the balance of the bridge is destroyed, as indicated by the current flowing in the bridge indicator (galvanometer). The sensitivity of this arrangement is very great and its "drift" (casual variations in bias, contact potential, plate voltage, etc.) is much less than that of the single-tube direct-coupled amplifier. The plate voltage, it will be noted, is applied equally to both tubes, so that any variations from this source are balanced out. The Wynn-Williams amplifier, in one form or another, has been used in medical and biological research for measuring the minute and slowly varying potentials generated by the tissues of living organisms.

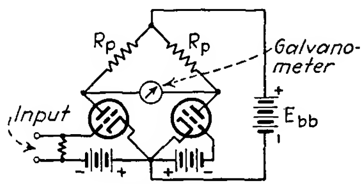


FIG. 212.—Wynn-Williams bridge amplifier.

When used as a current amplifier, rather than as a voltage amplifier, it is customary to employ electrometer tubes in the Wynn-Williams circuit, since the grid current then does not subtract from, or add to, the measured current in the grid resistor.

76. Illumination- and Color-measurement Circuits.—Measurement of light and color (photometry, densitometry, spectrophotometry) by photoelectric means has grown rapidly because of the many advantages which phototubes and photocells possess over the human eye. Among these advantages are speed, lack of fatigue, full sensitivity in the end regions of the spectrum, and automatic operation. The types of device used vary from a simple portable exposure meter used in photography, consisting of a self-generating photocell and a meter, to a fully automatic recording spectrophotometer occupying a table 10 by 4 ft.

Since measurement, rather than relaying, is of first importance in the purpose of these instruments, the phototube does not enjoy much advantage over the self-generating cell, which can operate a meter directly without auxiliary equipment. Consequently many of the simpler photometers and color comparators employ the self-generating cell.

Single-tube Illumination Meters, Photometers, and Densitometers.—A single self-generating photocell and a meter, or a

phototube, amplifier, and meter are widely used for illumination measurement, substitution photometry and densitometry. The output meter is calibrated directly in foot-candles or some other unit representing the illumination value, for an average white-light spectral distribution. The indications must, of course, be corrected if colored light is used. For photometry, the photocell or phototube views light alternately from a standard source and from the source under test, the ratio of the two illuminations, together with the geometric proportions of the test setup, determining the relative light output of the standard and sample. Since only one tube is used, it must be moved from the one position to the other in making the comparison.

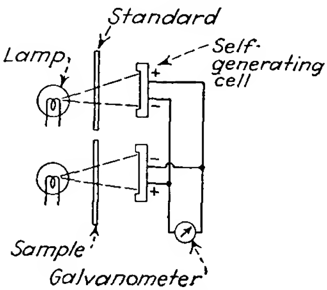


FIG. 213.—Transparency-comparison-meter arrangement.

For densitometry, comparing the light transmission or opacity of paper, textiles, photographic negatives, etc., a source of light is set up before the tube, and the meter reading is compared when the transmission sample is in front of the tube and when the tube views the light directly. The ratio of the two readings is a measure of the transmission properties of the sample.

Two-tube Units.—Fully automatic operation in photometry and densitometry is obtainable if two tubes or cells are used, since there is then no necessity of the operator's moving the tube from the sample to the standard. One such arrangement is shown in Fig. 213. In the two-tube units, periodic checks must be made to see that the two tubes and amplifiers possess the same sensitivity, *i.e.*, that no relative drift in characteristics has occurred. One highly sensitive type of a two-tube comparator is shown in Fig. 214. The two phototubes are connected in series, anode to cathode, and fed from a high-voltage source of good regulation. If the light on the two tubes is the same, the applied voltage distributes itself equally between the two tubes.

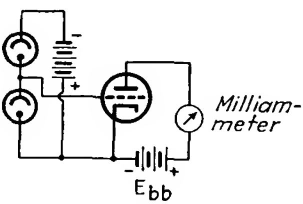


FIG. 214.—Double-phototube light-comparison circuit.

However, if one tube receives more light than the other, its resistance decreases and the voltage distribution between the two tubes changes. Since the tubes are operated at saturation, their current-voltage characteristics are very flat, and a small change in light results in a very large change in the voltage distribution. If a low-grid-current vacuum-grid voltmeter is used to measure the voltage across one of the phototubes, its plate current will undergo large changes as the relative light on the two phototubes changes. A sensitivity of better than 0.1 per cent can be achieved readily. The device measures the difference in light falling on the two tubes only; it does not reveal the absolute magnitude of the light itself. Another useful two-tube arrangement is the bridge-type photometer, in which the two tubes are placed as arms in a Wheatstone bridge. Any relative change in light unbalances the bridge and causes displacement of the balance indicator (galvanometer). A single phototube is often combined with a Wynn-Williams amplifier for high-precision work.

Color Comparators.—Color comparisons of dyes, both in solution and on papers and textiles, of printing inks, of various solutions in industrial processing and medical diagnosis, etc., may be carried out in various photoelectric instruments. The simplest is a single-tube illumination meter, fitted with a filter holder. Three filters are usually used, covering the visible spectrum in three slightly overlapping regions. The meter reading is taken eight times, four times with the sample and four times with the standard, with white light and with each of the three filters in place. If the readings for the sample agree with those of the standard, then a complete color match is indicated. Any variation between the readings indicates the portion of the spectrum in which the lack of match occurs. In using this method it is important that the tube view both sample and standard from the same point of view, otherwise apparent changes due to the reflection angle may be introduced. Ordinarily any change in phototube or photocell characteristics is balanced out in making the measurement since the same tube views both sample and standard. If a two-tube unit is used, to speed the comparison process, it is advisable to cross-check (test both tubes with the standard in position) before beginning each run.

Work with Low Light Levels.—In light-measurement work it is sometimes necessary to work with a source of light which is so weak that a reliable reading cannot be obtained from the phototube and its associated amplifier tube. If additional amplification is attempted with a direct-coupled amplifier, the variation in the contact potentials, etc., makes the measurement inaccurate at best. A simple means of avoiding the problem is the use of a light chopper, a rotating desk between light source and phototube containing several holes. The light entering the phototubes is thus interrupted periodically at an audio-frequency rate, and its output current contains an alternating-current component which can be amplified accurately in a multistage alternating-current amplifier. Another method of introducing an alternating-current component involves the use of the first amplifier tube as a nonlinear modulator, as shown in Fig. 215. An alternating-current signal (60-cycle alternating current is usually used) is introduced in series with the grid circuit and the direct phototube voltage. As the phototube current changes, the changing bias on the tube allows more or less of the alternating-current signal to appear in the plate circuit, where it is passed on to succeeding amplifier stages. The direct-current component is eliminated by the coupling circuit between the modulator tube and the amplifier following it.

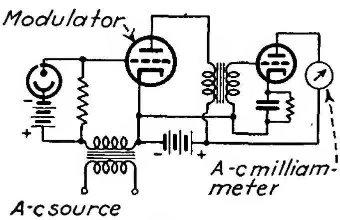


FIG. 215.—Modulator arrangement for introducing alternating-current component.

Automatic Recording Spectrophotometry.—The most elaborate of all electronic color-measuring devices is the automatic recording spectrophotometer, developed by Hardy. It is used for making a complete, automatic record of the spectral energy composition of any colored sample, either by transmitted or by reflected light. The machine covers the range from 4000 to 7000Å. in about 3 min., plotting the curve of the distribution as it goes. Remarkably enough the device uses one phototube only. Elaborate precautions have been incorporated in the design of the instrument to prevent changes in voltage or in the spectral characteristics of the phototube from affecting the accuracy of the measurement.

The diagram of the spectrophotometer is shown in Fig. 216. The light source, a small lamp, produces a narrow parallel beam of white light which passes first through a prism, which breaks the beam up into a spectrum. The spectrum is directed toward a mirror which reflects a portion of the spectrum to a second prism, which still further disperses the colors. The light emerging from the second prism is thus highly monochromatic. By rotating the mirror between the two prisms different portions of the first spectrum are presented to the second prism. Accordingly any color in the spectrum may be obtained from the

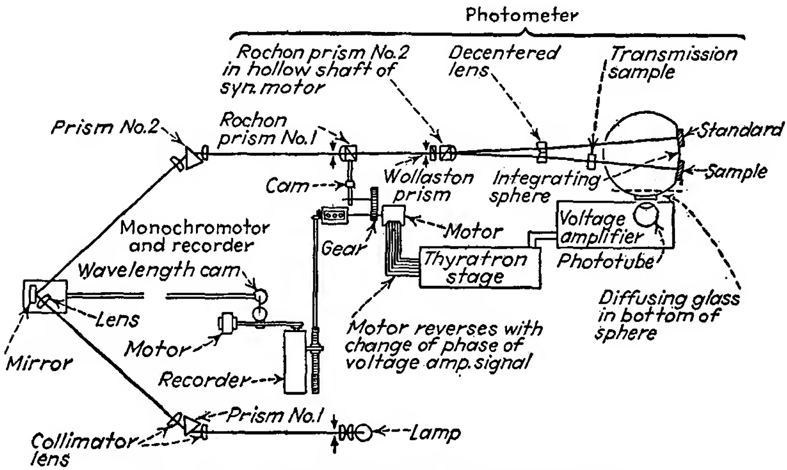


FIG. 216.—Automatic recording spectrophotometer.

second prism in a high degree of purity. This pure-colored light travels through a Rochon prism, which polarizes the light (causes the light vibrations to occur in one plane only). The polarized light then enters a second Rochon prism, which is rotated about the axis of the light beam by a motor, at a fixed rate of speed. When the optical axes of the two Rochon prisms are at right angles, no light passes the second Rochon prism, hence the output of the second Rochon prism is an alternating light of approximately pure sine-wave form. This "alternating-current light beam" enters an integrating sphere where it falls alternately on the sample and standard (the standard is usually a pure-white surface of magnesium carbonate). If both the sample and the standard have the same reflection coefficient

for the color of light present, the light flux within the integrating sphere remains constant, and the phototube which views its interior receives a direct-current component only. If, however, the sample has more or less reflectance than the sample, then an alternating-current component exists in the light flux, and this component is reproduced in the phototube current. The phototube current feeds an alternating-current amplifier, which responds therefore only when the reflection coefficients are not the same. The output of this amplifier, through a thyratron stage, drives an alternating-current motor. This alternating-current motor is coupling the first Rochon prism, and rotates it through a sufficient angle to redistribute the light between sample and standard, so that the alternating-current component disappears, whereupon the motor rotation ceases. As the spectrum mirror swings, producing light of different colors, the relative reflection between sample and standard is thus continuously measured for each spectrum color. The rotation of the correcting motor, being a measure of the relative reflectance, is used to drive a recording pencil over a sheet of plotting paper which is wrapped on a cylinder. This cylinder is in turn coupled to the swinging spectrum mirror, so that the ordinate on the graph paper which is presented to the recording pencil is always proportional to the wave length of the color light then passing through the instrument. Consequently a complete spectral distribution curve is plotted in the paper automatically.

Since only one phototube is used, and since it receives light of one color only at a time, measuring the relative reflectance at that wave length only, spectral response is of no consequence so long as the phototube has some sensitivity to all the wave lengths in the spectrum. Consequently changes in its response do not affect the measurement. Since the correction amplifier and motor operate on alternating current, direct-current drift cannot affect the level. Furthermore, the gain of the amplifier may change, the only effect being that the motor will take slightly more or less time to perform the necessary correction. The device is thus as nearly foolproof as any machine of its complexity could hope to be. It has already found wide use in the paint, dye, pigment, and ink industries, where accurate and unchanging records of color composition are essential for matching and reproduction purposes.

Problems

1. A tube (characteristics of Fig. 152) is connected in the relay circuit of Fig. 198. The relay has a coil whose resistance is 2000 ohms. The relay opens when the coil current is less than 4 ma. The applied plate battery is 250 volts. Determine the size of the grid battery necessary to insure the relay's opening when the grid circuit contact is closed.

2. Calculate the time at which the relay will open after the charging contacts are closed in the time-delay circuit shown in Fig. 199. Constants as in Prob. 1. $R = 1{,}600{,}000$ ohms $C = 5.0\mu$f. $E_{cc} = -60$ volts.

3. In a light-quantity meter a phototube is used to charge a condenser in the grid circuit of the relay tube in Prob. 2. Calculate the time at which the relay will open after a light beam of 2 lumens falls on the phototube. The battery in the grid circuit is large enough to saturate the phototube at all times (*i.e.*, the condenser is charged with a constant current). The phototube sensitivity is 25μa. per lumen.

4. Calculate the value of R_c in the voltage regulation circuit (Fig. 208) which will produce proper regulation when $R_1 = 5000$ ohms, $R_2 = 15{,}000$ ohms, and the mutual conductance of the tube is 1500 micromhos. The maximum and minimum values of plate current in the tube are 1 and 15 ma., respectively. Over what range may the input-voltage change before the regulation ceases, assuming the g_m of the tube does not change?

5. Compute and plot the calibration (alternating-current amplitude of input-voltage *versus* direct-current component of output current) of the diode rectifier voltmeter shown in Fig. 210*a*, by graphical means, using the characteristics of Fig. 57.

Bibliography

HENNEY, KEITH: "Electron Tubes in Industry," McGraw-Hill Book Company, Inc., New York, 1937.

GULLIKSEN and VEDDER: "Industrial Electronics," John Wiley & Sons, Inc., New York, 1936.

CHAMBERS, D. E.: Applications of Electron Tubes in Industry, *Elec. Eng.*, January, 1935.

MARTIN, S. R.: Thyratron Control for Resistance Welding Machines, *Welding*, May and June, 1932.

KING, W. R.: Photoelectric Relays, *Gen. Elec. Rev.*, August, 1932.

WALKER and LANCE: "Photoelectric Cell Applications," Pitman, 1935.

"TUBES At Work," a regular department of *Electronics*, September, 1935, to date.

MACARTHUR, E. D.: "Electronics and Electron Tubes," John Wiley & Sons, Inc., New York, 1936, Chap. VIII.

MORECROFT, J. H.: "Electron Tubes and Their Applications," John Wiley & Sons, Inc., New York, 1934.

ZWORYKIN and WILSON: "Photocells and Their Application," John Wiley & Sons, Inc., New York, 1934.

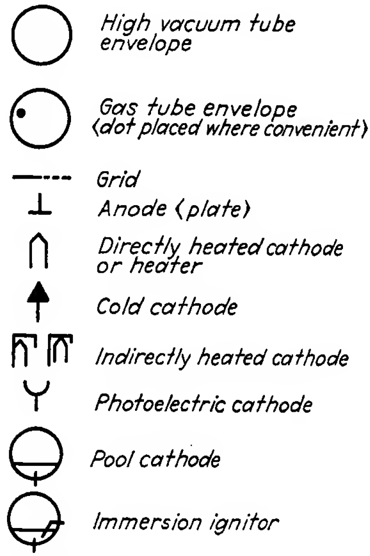


FIG. 217.—Institute of Radio Engineers' standard diagram symbols for electron tubes.

APPENDIX I

ELECTRON-TUBE LETTER SYMBOLS

Quantity	Symbol
Instantaneous total grid voltage	e_c
Instantaneous total plate voltage	e_b
Instantaneous total grid current	i_c
Instantaneous total plate current	i_b
Average or quiescent value of grid voltage	E_c
Average or quiescent value of plate voltage	E_b
Average or quiescent value of grid current	I_c
Average or quiescent value of plate current	I_b
Instantaneous value of varying component of grid voltage	e_g
Instantaneous value of varying component of plate voltage	e_p
Instantaneous value of varying component of grid current	i_g
Instantaneous value of varying component of plate current	i_p
Effective or maximum value of varying component of grid voltage	E_g
Effective or maximum value of varying component of plate voltage	E_p
Effective or maximum value of varying component of grid current	I_g
Effective or maximum value of varying component of plate current	I_p
Filament or heater terminal voltage	E_f
Grid battery voltage	E_{cc}
Plate battery voltage	E_{bb}
Filament or heater current	I_f
Total electron emission	I_s
Conductance of electrode j	g_j
Resistance of electrode j	r_j
Plate conductance	g_p
Plate resistance	r_p
Grid conductance	g_g
Grid resistance	r_g
Transconductance from electrode k to electrode j	g_{jk}
Grid-plate transconductance (mutual conductance)	g_m
Mu-factor, electrodes j and k	$\mu_{jkl} = \left(\dfrac{\delta e_j}{\delta e_k}\right)_{i_l \text{ constant}}$

Quantity	Symbol
Amplification factor	$\mu = -\left(\frac{\delta e_p}{\delta e_g}\right)_{i_p \text{ constant}}$
Grid-plate capacitance	C_{gp}
Grid-cathode capacitance	C_{gk}
Plate-cathode capacitance	C_{pk}
Grid-heater capacitance	C_{gh}
Plate-heater capacitance	C_{ph}
Grid capacitance	C_g
Plate capacitance	C_p
Cathode capacitance	C_k
Tungsten sensitivity	S_T
2870 tungsten sensitivity	S_{2870}
Radiant flux	Φ, ϕ
Luminous flux	F, f

APPENDIX II

DEFINITIONS OF ELECTRONIC TERMS

Compiled by committees under the sponsorship of the American Institute of Electrical Engineers as part of the program of the American Standards Association

Author's Note.—This listing of definitions is taken verbatim from "Group 74—Electronics," which is a division of the "Revised Report on Proposed American Standard Definitions of Electrical Terms." The definitions are the result of the work of two subcommittees of the A.I.E.E., the "Sectional Committee No. C60 on Vacuum Tubes for Industrial Purposes," and "Subcommittee No. 13A—Electronics." The definitions have been offered for approval by the Sectional Committee on Electrical Definitions but no final action has been taken at the time of writing (April, 1938). While not yet official American standards, they are the most authoritative definitions available at the present time. For further information, the reader may address the American Institute of Electrical Engineers, 33 West 39th St., New York, N.Y.

Section 05. General

Electronics. 74.05.005. Electronics is that branch of science and technology which relates to the conduction of electricity through gases or in vacuo.

Electron. 05.10.040. An electron is the natural, elementary quantity of negative electricity.

The quantity of electricity on an electron is 1.592×10^{-19} coulomb, or 4.774×10^{-10} electrostatic units.*

The mass of an electron at rest is 9.00×10^{-28} gram.*

Ion. 05.10.050. An ion is an electrified portion of matter of subatomic, atomic, or molecular dimensions.

Electrode. 60.05.030. An electrode is a conductor belonging to the class of metallic conductors, but not necessarily a metal, through which a current enters or leaves an electrolytic cell, arc, furnace, vacuum tube, gaseous discharge tube, or any conductor of the non-metallic class.

Electron Emission. 74.05.025. Electron emission is the liberation of electrons from an electrode into the surrounding space. Quantitatively, it is the rate at which electrons are emitted from an electrode.

Thermionic Emission. 74.05.030. Thermionic emission is electron or ion emission due directly to the temperature of the emitter.

* See text, p. 57.

Secondary Emission. 74.05.035. Secondary emission is electron emission due directly to the impact of electrons or ions.

Ionization. 05.10.055. Ionization is the process of producing ions.

Ionization Current. 75.05.080. Ionization current is the electric current resulting from the movement of electric charges produced by the action of an applied electric field upon an ionized medium.

Gas Current. 74.05.050. A gas current is a current flowing to an electrode and composed of positive ions which have been produced as a result of gas ionization by an electron current flowing between other electrodes.

Section 10. Vacuum Tubes

Vacuum Tube. 74.10.005. (Electron Tube). A vacuum tube is a device consisting of an evacuated enclosure containing a number of electrodes between two or more of which conduction of electricity through the vacuum or contained gas may take place.

High-vacuum Tube. 74.10.010. A high-vacuum tube is a vacuum tube evacuated to such a degree that its electrical characteristics are essentially unaffected by gaseous ionization.

Thermionic Tube. 74.10.015. A thermionic tube is a vacuum tube in which one of the electrodes is heated for the purpose of causing electron or ion emission from that electrode.

Gas Tube. 74.10.020. A gas tube is a vacuum tube in which the pressure of the contained gas or vapor is such as to affect substantially the electrical characteristics of the tube.

Mercury-vapor Tube. 74.10.025. A mercury-vapor tube is a gas tube in which the active contained gas is mercury vapor.

Phototube. 74.10.030. (Photoelectric Tube). A phototube is a vacuum tube in which one of the electrodes is irradiated for the purpose of causing electron emission from that electrode.

Cathode-ray Oscillograph Tube. 74.10.035. A cathode-ray oscillograph tube is a vacuum tube in which the deflection of an electron beam, effected by means of applied electric and/or magnetic fields, indicates the instantaneous values of the actuating voltages and/or currents.

Diode. 74.10.040. A diode is a two-electrode vacuum tube containing an anode and a cathode.

Triode. 74.10.045. A triode is a three-electrode vacuum tube containing an anode, a cathode, and a control electrode.

Tetrode. 74.10.050. A tetrode is a four-electrode vacuum tube containing an anode, a cathode, a control electrode, and one additional electrode ordinarily in the nature of a grid.

Pentode. 74.10.055. A pentode is a five-electrode vacuum tube containing an anode, a cathode, a control electrode, and two additional electrodes ordinarily in the nature of grids.

Hexode. 74.10.060. A hexode is a six-electrode vacuum tube containing an anode, a cathode, a control electrode, and three additional electrodes ordinarily in the nature of grids.

Heptode. 74.10.065. A heptode is a seven-electrode vacuum tube containing an anode, a cathode, a control electrode, and four additional electrodes ordinarily in the nature of grids.

Octode. 74.10.070. An octode is an eight-electrode vacuum tube containing an anode, a cathode, a control electrode, and five additional electrodes ordinarily in the nature of grids.

Multiple-unit Tube. 74.10.075. A multiple-unit tube is a vacuum tube containing within one envelope two or more groups of electrodes associated with independent electron streams.

NOTE.—A multiple-unit tube may be so indicated, as for example: duodiode, duotriode, diode-pentode, duodiode-triode, duodiode-pentode, and triode-pentode.

Section 15. Vacuum-tube Electrodes

Anode of a Vacuum Tube. 74.15.005. (Plate.) An anode of a vacuum tube is an electrode to which a principal electron stream flows.

Plate. 74.15.010. Plate is a common name for the principal anode in a vacuum tube.

Cathode of a Vacuum Tube. 74.15.015. A cathode of a vacuum tube is an electrode which is the primary source of an electron stream.

Indirectly Heated Cathode. 74.15.020. (Equipotential Cathode, Unipotential Cathode.) An indirectly heated cathode is a cathode of a thermionic tube to which heat is supplied by an independent heater element.

Heater. 74.15.025. A heater is an electric heating element for supplying heat to an indirectly heated cathode.

Filament. 74.15.030. A filament is a cathode of a thermionic tube, usually in the form of a wire or ribbon, to which heat may be supplied by passing current through it.

Control Electrode. 74.15.035. A control electrode is an electrode on which a voltage is impressed to vary the current flowing between two or more other electrodes.

Grid. 74.15.040. A grid is an electrode having one or more openings for the passage of electrons or ions.

Control Grid. 74.15.045. A control grid is a grid, ordinarily placed between the cathode and an anode, for use as a control electrode.

Screen Grid. 74.15.050. A screen grid is a grid placed between a control grid and an anode, and usually maintained at a fixed positive potential, for the purpose of reducing the electrostatic influence of the anode in the space between the screen grid and the cathode.

Space-charge Grid. 74.15.055. A space-charge grid is a grid which is placed adjacent to the cathode and positively biased so as to reduce the limiting effect of space charge on the current through the tube.

Suppressor Grid. 74.15.060. A suppressor grid is a grid which is interposed between two electrodes (usually the screen grid and plate), both positive with respect to the cathode, in order to prevent the passing of secondary electrons from one to the other.

Section 20. Electrode Voltage, Current, and Power

Electrode Voltage. 74.20.005. Electrode voltage is the voltage between an electrode and a specified point of the cathode.

Anode Voltage. See electrode voltage, **74.20.005.**

Peak (or Crest) Forward Anode Voltage. 74.20.015. Peak (or crest) forward anode voltage is the maximum instantaneous anode voltage in the direction in which the tube is designed to pass current.

Peak (or Crest) Inverse Anode Voltage. 74.20.020. Peak (or crest) inverse anode voltage is the maximum instantaneous anode voltage in the direction opposite to that in which the tube is designed to pass current.

Heater Voltage. 74.20.025. Heater voltage is the voltage between the terminals of a heater.

Filament Voltage. 74.20.030. Filament voltage is the voltage between the terminals of a filament.

Grid Voltage. See electrode voltage, **74.20.005.**

Direct Grid Bias. 74.20.040. Direct grid bias is the direct component of grid voltage.

NOTE.—This is commonly called grid bias.

Electrode Current. 74.20.045. Electrode current is the current passing to or from an electrode through the vacuous space.

Anode Current. See electrode current, **74.20.045.**

Cathode Current. 74.20.055. Cathode current is the total current passing to or from the cathode through the vacuous space.

Heater Current. 74.20.060. Heater current is the current flowing through a heater.

Filament Current. 74.20.065. Filament current is the current supplied to a filament to heat it.

Grid Current. See electrode current, **74.20.045.**

Grid Emission. 74.20.075. Grid emission is electron or ion emission from a grid.

Leakage Current. 74.20.080. A leakage current is a conductive current which flows between two or more electrodes by any path other than across the vacuous space.

Electrode Characteristic. 74.20.085. An electrode characteristic is a relation, usually shown by a graph, between an electrode voltage and current, other electrode voltages being maintained constant.

Transfer Characteristic. 74.20.090. A transfer characteristic is a relation, usually shown by a graph, between the voltage of one electrode and the current to another electrode, all other voltages being maintained constant.

Emission Characteristic. 74.20.095. An emission characteristic is a relation, usually shown by a graph, between the emission and a factor controlling the emission (as temperature, voltage, or current of the filament or heater).

Electrode Dissipation. 74.20.100. Electrode dissipation is the power dissipated in the form of heat by an electrode as a result of electron and/or ion bombardment.

Grid Driving Power. 74.20.105. Grid driving power is the average product of the instantaneous value of the grid current and the alternating component of the grid voltage over a complete cycle.

NOTE.—This comprises the power supplied to the biasing device and the grid dissipation.

Section 25. Electrode Impedances and Admittances

Mu Factor. 74.25.005. The mu factor is the ratio of the change in one electrode voltage to the change in another electrode voltage, under the conditions that a specified current remains unchanged and that all other electrode voltages are maintained constant. It is a measure of the relative effect of the voltages of two electrodes on the current in the circuit of any specified electrode.

NOTE.—As most precisely used, the term refers to infinitesimal changes.

Amplification Factor. 74.25.010. The amplification factor is the ratio of the change in plate voltage to a change in control electrode voltage, under the conditions that the plate current remains unchanged and that all other electrode voltages are maintained constant. It is a measure of the effectiveness of the control-electrode voltage relative to that of the plate voltage on the plate current. The sense is usually taken as positive when the voltages are changed in opposite directions.

NOTE.—As most precisely used, the term refers to infinitesimal changes. Amplification factor is a special case of mu factor.

Electrode Admittance. 74.25.015. Electrode admittance is the quotient of the alternating component of the electrode current by the alternating component of the electrode voltage, all other electrode voltages being maintained constant.

NOTE.—As most precisely used, the term refers to infinitesimal amplitudes.

Transadmittance. 74.25.020. Transadmittance from one electrode to another is the quotient of the alternating component of the current of the second electrode by the alternating component of the voltage of the first electrode, all other electrode voltages being maintained constant.

NOTE.—As most precisely used, the term refers to infinitesimal amplitudes.

Electrode Impedance. 74.25.025. Electrode impedance is the reciprocal of the electrode admittance.

Electrode Conductance. 74.25.030. Electrode conductance is the quotient of the electrode alternating current by the in-phase component of the electrode alternating voltage, all other electrode voltages being maintained constant.

NOTE.—This is a variational and not a total conductance. As most precisely used, the term refers to infinitesimal amplitudes.

Plate Conductance. See electrode conductance, **74.25.030.**

Grid Conductance. See electrode conductance, **74.25.030.**

Transconductance. 74.25.045. Transconductance from one electrode to another is the quotient of the alternating current of the second electrode by the in-phase component of the alternating voltage of the first electrode, all other electrode voltages being maintained constant.

NOTE.—As most precisely used, the term refers to infinitesimal amplitudes. See control-grid-plate transconductance.

Control-grid-plate Transconductance. 74.25.050. (Transconductance, Mutual Conductance.) Control-grid-plate transconductance is the name for the plate-current to control-grid voltage transconductance.

NOTE.—This is ordinarily the most important transconductance and is commonly understood when the term transconductance or mutual conductance is used.

Mutual Conductance. See control-grid-plate transconductance, **74.25.050.**

Electrode Resistance. 74.25.060. Electrode resistance is the reciprocal of the electrode conductance.

NOTE.—This is the effective parallel resistance and is not the real component of the electrode impedance.

Plate Resistance. See electrode resistance, **74.25.060.**

Capacitance. 74.25.070. The various capacitances of any system of conductors are defined in definitions 05.15.060 to 05.15.092 inclusive, which are directly applicable to the case of the vacuum tube.

Capacitance between Two Conductors. 05.15.060. The capacitance between two conductors is equal to the ratio of the charge placed on either conductor to the resulting change in potential difference, provided the conductors have received equal and opposite charges. Ignored conductors, either concealed or intentionally neglected, may be present, but their charges must not be changed during the measurement.

Interelectrode Capacitance. 74.25.075. Interelectrode capacitance is the direct capacitance between two electrodes.

Electrode Capacitance. 74.25.080. Electrode capacitance is the capacitance of one electrode to all other electrodes connected together.

Input Capacitance. 74.25.085. The input capacitance of a vacuum tube is the sum of the direct capacitances between the control grid and the cathode and such other electrodes as are operated at the alternating potential of the cathode.

NOTE.—This is not the effective input capacity which is a function of the impedances of the associated circuit.

Output Capacitance. 74.25.090. The output capacitance of a vacuum tube is the sum of the direct capacitances between the output electrode (usually the plate) and the cathode and such other electrodes as are operated at the alternating potential of the cathode.

NOTE.—This is not the effective output capacity which is a function of the impedances of the associated circuit.

Section 30. Amplifiers

Amplifier. 05.45.155. An amplifier is a device for increasing the energy associated with any phenomenon without appreciably altering its quality.

Class A Amplifier. 74.30.010. A class A amplifier is an amplifier in which the grid bias and alternating grid voltages are such that plate current in a specific tube flows at all times.

NOTE.—To denote that grid current does not flow during any part of the input cycle, the suffix 1 may be added to the letter or letters of the class identification. The suffix 2 may be used to denote that grid current flows during some part of the cycle.

Class AB Amplifier. 74.30.015. A class AB amplifier is an amplifier in which the grid bias and alternating grid voltages are such that plate

current in a specific tube flows for appreciably more than half but less than the entire electrical cycle.

NOTE.—See note under Class A Amplifier.

Class B Amplifier. 74.30.020. A class B amplifier is an amplifier in which the grid bias is approximately equal to the cut-off value so that the plate current is approximately zero when no exciting grid voltage is applied, and so that plate current in a specific tube flows for approximately one-half of each cycle when an alternating grid voltage is applied.

NOTE.—See note under Class A Amplifier.

Class C Amplifier. 74.30.025. A class C amplifier is an amplifier in which the grid bias is appreciably greater than the cut-off value so that the plate current in each tube is zero when no alternating grid voltage is applied, and so that plate current flows in a specific tube for appreciably less than one-half of each cycle when an alternating grid voltage is applied.

NOTE.—See note under Class A Amplifier.

Section 35. Gas Tubes

Tube Voltage Drop. 74.35.005. Tube voltage drop in a gas tube is the anode voltage during the conducting period.

Critical Grid Voltage. 74.35.010. Critical grid voltage in a gas tube is the instantaneous value of grid voltage when the anode current starts to flow.

Critical Grid Current. 74.35.015. Critical grid current in a gas tube is the instantaneous value of grid current when the anode current starts to flow.

Control Characteristic. 74.35.020. The control characteristic of a gas tube is a relation, usually shown by a graph, between critical grid voltage and anode voltage.

Cathode Heating Time. 74.35.025. Cathode heating time is the time required for the cathode to attain operating temperature with normal voltage applied to the heating element.

Tube Heating Time. 74.35.030. Tube heating time in a mercury-vapor tube is the time required for the coolest portion of the tube to attain operating temperature.

Section 40. Phototubes

Static Sensitivity of a Phototube. 74.40.005. Static sensitivity of a phototube is the quotient of the direct anode current by the incident radiant flux of constant value.

Dynamic Sensitivity of a Phototube. 74.40.010. Dynamic sensitivity of a phototube is the quotient of the alternating component of anode current by the alternating component of incident radiant flux.

NOTE.—As most precisely used, the term refers to infinitesimal amplitudes. This is a variational and not a total sensitivity.

Luminous Sensitivity of a Phototube. 74.40.015. Luminous sensitivity of a phototube is the quotient of the anode current by the luminous flux.

2870 Tungsten Sensitivity of a Phototube. 74.40.020. 2870 tungsten sensitivity of a phototube is the quotient of the anode current by the total

incident luminous flux in lumens from a tungsten filament lamp at a color temperature of 2870 degrees Kelvin.

Gas Amplification Factor of a Phototube. 74.40.025. Gas amplification factor of a phototube is the factor of increase in the sensitivity of a gas phototube due solely to the ionization of the contained gas.

NOTE.—For a gas phototube having a structure such as to permit saturation to occur at a voltage less than that causing appreciable ionization, the gas amplification factor at a specified operating voltage is the ratio of the sensitivity measured at that voltage to the sensitivity measured at the saturation voltage.

Current Wave-length Characteristic of a Phototube. 74.40.030. Current wave-length characteristic of a phototube is a relation usually shown by a graph between the direct anode current per unit energy of the incident radiant flux and the wave length of the flux.

Section 45. Cathode-ray Oscillograph Tubes

Gas Focusing. 74.45.005. Gas focusing is a method of focusing an electron stream in which focus is produced through the action of ionized gas.

Electrostatic Focusing. 74.45.010. Electrostatic focusing is a method of focusing an electron stream in which focus is produced through the action of an electric field.

Magnetic Focusing. 74.45.015. Magnetic focusing is a method of focusing an electron stream in which focus is produced through the action of a magnetic field.

Deflection Sensitivity of a Cathode-ray Oscillograph Tube. 74.45.020. The deflection sensitivity of a cathode-ray oscillograph tube is the quotient of the displacement of the electron beam at the place of impact by the change in the deflecting field.

NOTE.—It is usually expressed in millimeters per volt applied between the deflecting plates or in millimeters per gauss of the deflecting magnetic field.

Deflection Factor of a Cathode-ray Oscillograph Tube. 74.45.025. The deflection factor of a cathode-ray oscillograph tube is the reciprocal of the deflection sensitivity.

Section 50. X-ray Tubes

X-ray Tube. 74.50.010. An X-ray tube is a vacuum tube designed for producing X-rays by accelerating electrons to a high velocity by means of an electrostatic field and then suddenly stopping them by collision with a target.

Gas X-ray Tube. 74.50.015. A gas X-ray tube is an X-ray tube in which the emission of electrons from the cathode is produced by positive ion bombardment.

Hot-cathode X-ray Tube. 74.50.020. A hot-cathode X-ray tube is a high-vacuum X-ray tube in which the cathode is heated in order to produce the emission of electrons.

Target. 74.50.025. (Anode; Anticathode).[1] A target is an electrode, or part of an electrode, on which cathode rays are focused and from which X-rays are emitted.

Section 95. Not Otherwise Classified

Rectifier. 15.50.010. A rectifier is a device which converts alternating current into uni-directional current by virtue of a characteristic permitting appreciable flow of current in one direction only. The operation of the device as a rectifier depends on its ability to offer very low resistance to current flow in one direction, but to offer very high or practically infinite resistance to flow of current in the opposite direction.

Mercury-arc Rectifier. 15.50.020. A mercury-arc rectifier is a rectifier which makes use of the rectifying properties of a mercury arc between electrodes enclosed in a highly evacuated chamber.

Metal-tank Mercury-arc Rectifier. 15.50.030. A metal-tank mercury-arc rectifier is a mercury-arc rectifier with the anodes and mercury cathode enclosed in a metal container or chamber.

Controlled Mercury-arc Rectifier. 15.50.035. A controlled mercury-arc rectifier is a mercury-arc rectifier in which one or more electrodes are employed to control the starting of the discharge.

Discharge Tube. 74.95.025. A discharge tube is an evacuated enclosure containing a gas at low pressure which permits the passage of electricity through the gas upon application of sufficient voltage.

Note.—The tube is usually provided with metal electrodes, but one form permits an electrodeless discharge with induced voltage.

Geissler Tube. 74.95.030. A Geissler tube is a special form of discharge tube for showing the luminous effects of discharges through rarefied gases.

Note.—The density of gas is roughly one-thousandth of that of the atmosphere.

Cathode Glow. 74.95.035. The cathode glow is the luminous glow which covers the surface of the cathode in a discharge tube, between the cathode and the cathode dark space.

Cathode Dark Space. 74.95.040. (Crookes Dark Space.) The cathode dark space is the relatively non-luminous region in a discharge tube between the cathode glow and the negative glow.

Negative Glow. 74.95.045. The negative glow is the luminous glow in a discharge tube between the cathode dark space and the Faraday dark space.

Faraday Dark Space. 74.95.050. The Faraday dark space is the relatively non-luminous region in a discharge tube between the negative glow and the positive column.

Positive Column. 74.95.055. The positive column is the luminous glow, often striated, in a discharge tube between the Faraday dark space and the anode.

APPENDIX III

MEANINGS OF ABBREVIATIONS AND SYMBOLS USED IN THE TEXT

Abbreviations

Å. Angstrom unit (of length). 1Å. = 10^{-8} cm.
amp. ampere.
cm. centimeter. 1 cm. = 0.3937 in.
c.p.s. cycles per second (commonly "cycles").
db. decibel.
dyne 2.248×10^{-6} lb.
e.m.u. electromagnetic unit (may be applied to current, voltage, fields, etc.).
erg 10^{-7} watt-sec.
e.s.u. electrostatic unit (may be applied to current, voltage, charge, fields, etc.).
f. farad.
ft. foot, feet.
g. gram. 1 gram = 0.03527 ounce.
hr. hour.
in. inch.
kc. kilocycle per second. 1 kc. = 1000 c.p.s.
km. kilometer. 1 km. = 1000 m.
kw. kilowatt. 1 kw. = 1000 watts.
lb. pound.
m. meter. 1 m. = 3.281 ft.
ma. milliampere. 1 ma. = 10^{-3} amp.
min. minute.
mm. millimeter. 1 mm. = 10^{-3} m.
sec. second.
μ micron. 1 μ = 10^{-6} m.
μa. microampere. 1 μa. = 10^{-6} amp.
μf. microfarad. 1 μf = 10^{-6} f.
$\mu\mu$f. micromicrofarad. 1 $\mu\mu$f = 10^{-12} f.
μmho micromho. 1 μmho = 10^{-6} amp. per volt.
/ (shilling mark) should be read "divided by" or "per."

Symbols

A thermionic constant.
b_0 thermionic constant.

C capacitance.
°C. degrees centigrade.
cos cosine.
d distance.
e, E voltage.
e electron charge.
$\epsilon = 2.718+$.
f frequency.
g gravitational constant.
g_m grid-plate transconductance, mutual conductance.
h Planck's constant $= 6.54 \times 10^{-27}$ erg-sec.
H magnetic field strength.
i, I current.
j imaginary operator $= \sqrt{-1}$.
K.E. kinetic energy.
°K. degrees Kelvin (degrees absolute) = °C. + 273°.
L inductance.
l length.
$\log_\epsilon$ natural logarithm.
$\log_{10}$ common logarithm.
m mass.
n number.
p pressure.
ϕ_0 work function.
π (pi) = 3.1416
Q quantity of electricity.
r radius.
R resistance.
s separation (distance).
S sensitivity.
sin sine.
T temperature.
t time.
tan tangent.
μ amplification factor.
v velocity.
W energy or work.
X electric field strength.
X_L inductive reactance.
X_C capacitive reactance.
Z impedance.

APPENDIX IV

ANSWERS TO PROBLEMS

Chapter II (pages 31, 32)

1. 1.10×10^{18} cm./sec./sec.

2. 0.32 cm.

3. 8.4×10^{8} cm./sec.
7.14×10^{-10} sec.

4. 3.21×10^{-10} ergs.
The battery.
The collecting electrode and the external circuit resistance, if any.

5. 4.2×10^{8} cm./sec.
2.24×10^{-10} amp.
4.46×10^{9} electrons.

6. 1.02×10^{25} electrons.
0.93×10^{-2} g.
62 per cent.
1.6 yr.

7. 1,130,000 volts.

Chapter III (pages 50, 51)

1. 3.75 volts.
5.12×10^{-12}ergs.

2. 7.2×10^{-12} ergs.
3.2×10^{-8} cm.

3. 2.4×10^{-12} ergs.
0.73×10^{8} cm./sec.

4. 0.345 amp.

5. 152 sq. cm.

6. 6×10^{14} c.p.s.
5.2×10^{14} c.p.s.
1.53×10^{15} c.p.s.

7. 9 per cent.
25 electrons for every original electron.
9,765,625 for every original electron.

Chapter IV (pages 70, 71)

1. 1.025×10^{9} cm./sec.
41 per cent.

2. 9.36 ma.
131 ma.

3. Assuming $\beta^2 = 1$,
$r/l = 14.68$.
$r_0/l = 1.35$.
E.g. for $l = 1$
$r = 14.68$ cm.
$r_0 = 1.35$ cm.

4. 6.5 volts.

5. 2.2 cm.

6. 1.76×10^{10} cm./sec. (neglecting Einstein-Lorenz increase in mass).

7. 4.43 e.m.u.

Chapter V (page 89)

1. $W_1 = -2.23 \times 10^{-11}$ ergs. $W_2 = -5.58 \times 10^{-12}$ ergs.
   $W_3 = -2.48 \times 10^{-12}$ ergs.
   $W_4 = -1.40 \times 10^{-12}$ ergs. $W_5 = -8.93 \times 10^{-13}$ ergs.
2. $r_1 = 0.517 \times 10^{-8}$ cm. $r_2 = 2.07 \times 10^{-8}$ cm. $r_3 = 4.65 \times 10^{-8}$ cm.
   $r_4 = 8.27 \times 10^{-8}$ cm. $r_5 = 12.9 \times 10^{-8}$ cm.
3. Hydrogen (single proton) 1.92×10^{14} cm./sec./sec.
   Electron 3.52×10^{17} cm./sec./sec.
   Mercury (singly charged ion) 0.96×10^{12} cm./sec./sec.
4. 3.33×10^{-5} cm.
   5.9×10^{-23} atmosphere.
5. 3.58 cm.
6. 0.04 amp. (electron current).
   0.00094 amp. (ion current).
7. Yes. 10 ions per electron.

Chapter VI (pages 131, 132)

1. 59 ma./watt.
2.

e_b	i_b
50 volts	10.6 ma.
100 volts	30.0 ma.
150 volts	55.0 ma.
200 volts	84.4 ma.
560 volts	396. ma.

3. 16 ma.
4. 2500 ohms.
5. Amplification:
   6C5 18.2
   24A 93.0
   6D6 133.0
   Load resistance:
   6C5 20,000 ohms.
   24A 600,000 ohms.
   6D6 800,000 ohms.
6. $\mu = 25$.
   $r_p = 16{,}660$ ohms.
   $R_g = 0.145$ cm.
   $R_p = 0.445$ cm.
   $l = 2.8$ cm.
7. $\mu = 19$.
   $g_m = 1520$ μmhos.
   $r_p = 12{,}500$ ohms.

Chapter VII (page 158)

1. 45 ma.
2. 340 ma.
   180 ma.
3. 345 ma.
   340 ma.
4. 100 μa.

Chapter VIII (page 180)

1. (a) 9.77 ma.
   (b) 7.85 ma.
2. (a) 0.152 ma.
   (b) 0.122 ma.
   (a) 5.47 ma.
   (b) 4.41 ma.
3. (a) 3.45 ma.
   (b) 2.78 ma.
4. 9 μa./lumen.
5. 14 μa./lumen.
6. 6.5×10^{-6} watts.
   0.39 per cent.

Chapter IX (page 198)

1. Na: 4.8×10^{14}, 5×10^{14}, 5.2×10^{14}, 6×10^{14}.
   Hg: 5×10^{14}, 7.3×10^{14}. Ne: 4.3×10^{14}, 5.3×10^{14}.
2. 1.76 volts; 2.20 volts.
3. Neon is 11.5 per cent as efficient as sodium.
4. Na: 5.3 μa./lumen. Ne: 5.9 μa./lumen.
5. 10,000 to 1.
   8.5 volts.
   Roughly 10,000.
   2.53 volts.
   By exciting a sodium atom.

Chapter X (pages 220, 221)

3. 308,400 volts.
   (*a*) 3.3×10^{10} cm./sec. (in excess of speed of light, not physically possible).
   (*b*) 2.3×10^{10} cm./sec.
   1390 watt-seconds (joules) = 1.39×10^{10} ergs.
4. No. of elements = (pictures per second) (no. of lines)^2 ($1\frac{1}{3}$). 16, 8,200; 24, 18,400; 60, 115,000; 120, 460,000; 240, 1,840,000; 343, 4,700,000; 441, 7,750,000.
5. 0.0085 in. in diameter.
   0.067 g.
6. −40,000 ohms.
7. External displacement of 2.25 cm. changes current from 15 ma. to 0.15 ma.

Chapter XI (page 252)

1. 0.49 ma.
   0.3 volt.
2. 150 c.p.s.: $+j8400$ ohms; 0.84 volt.
   160 c.p.s.: $-j91{,}000$ ohms; 9.1 volts
   170 c.p.s.: $-j7600$ ohms; 0.76 volt.
3. 100 volts d.c., 10 volts a.c.: 15 ma. d.c., 3.5 ma. a.c.
   300 volts d.c., 30 volts a.c.: 63 ma. d.c., 16 ma. a.c.
4. $E_c = -4.0$ volts.
   $E_b = 115$ volts.
   $I_b = 2.5$ ma.
5. 0.017 watt.
6. 0.305 watt.
   5.5 per cent.
7. 0.022 watt.
   1.4 watts.
   1.5 per cent.
   Because of the larger values of E_b and I_b

Chapter XII (page 270)

1. 80 watts.
4. 17 cycles per second.

Chapter XIII (pages 304, 305)

1. 18.13 db, 23.98 db, −30 db, 62.43 db.
   24.08 db, −48.16 db, 80 db, −13.12 db.
2. 0.465 watt.
   58.5 per cent.
3. 100 c.p.s.: 5.7.
   1000 c.p.s.: 7.3.
   10,000 c.p.s.: 6.6.
4. 66.6 per cent.
   999,000 and 1,001,000 c.p.s.
5. Signal: 0.12 ma.
   Direct component: 0.235 ma.

Chapter XIV (page 331)

1. Any voltage more negative than −10 volts.
2. 20 sec.
3. 1 sec.
4. 2660 ohms.
   37.3 volts.

INDEX